

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 10, 2005, 07:52:23 ; Search time 311.571 Seconds
(without alignments)
4400.037 Million cell updates/sec

Title: US-10-600-070B-3

Perfect score: 3667
Sequence: 1 tgcctgcataagagagaat.....ctataacataagggctacaa 3667

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3392430 seqs, 186927314 residues

Total number of hits satisfying chosen parameters: 6784860

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA_New:*

1: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:.*
2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:.*
3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:.*
4: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:.*
5: /cgn2_6/ptodata/2/pubpna/BCT_NEW_PUB.seq:.*
6: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:.*
7: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:.*
8: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:.*
9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:.*
10: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	44.8	1.2	191684	US-11-121-086-2	Sequence 2, Appli
C 2	42.4	1.2	169495	US-11-121-086-61	Sequence 61, Appli
C 3	42	1.1	1199	US-10-750-185-40551	Sequence 40551, A
C 4	40.6	1.1	161874	US-11-121-086-75	Sequence 75, Appli
C 5	40.2	1.1	31028	US-10-829-826B-21	Sequence 21, Appli
C 6	40.2	1.1	31028	US-10-829-826B-22	Sequence 22, Appli
C 7	40.2	1.1	31028	US-10-829-826B-26	Sequence 26, Appli
C 8	40.2	1.1	31100	US-10-829-826B-24	Sequence 24, Appli
C 9	39.8	1.1	1450	US-11-140-417-3	Sequence 3, Appli
C 10	39.8	1.1	168516	US-11-121-086-3	Sequence 3, Appli
C 11	39.6	1.1	600	US-10-750-185-3575	Sequence 3575, Ap
C 12	39.4	1.1	7402	US-10-750-185-27313	Sequence 27313, A
C 13	39.4	1.1	199321	US-11-121-086-10	Sequence 10, Appli
C 14	39.2	1.1	1317	US-11-140-417-1	Sequence 1, Appli
C 15	39.2	1.1	11462	US-11-140-417-22	Sequence 22, Appli
C 16	39.2	1.1	120096	US-11-121-086-24	Sequence 24, Appli
C 17	39	1.1	913	US-10-750-185-61576	Sequence 61576, A
C 18	39	1.1	153376	US-11-121-086-5	Sequence 5, Appli
C 19	38.6	1.1	126552	US-11-121-086-1	Sequence 1, Appli
C 20	38.4	1.0	1082144	US-11-117-187-211	Sequence 211, App
C 21	38.2	1.0	1360	US-10-750-185-60979	Sequence 60979, A
C 22	38.2	1.0	7402	US-10-750-185-27313	Sequence 27313, A
C 23	38	1.0	15720	US-11-108-172-1058	Sequence 1058, Ap

24	38	1.0	176503	US-11-121-086-53	Sequence 53, Appli
25	37.8	1.0	2956	US-10-689-742-103	Sequence 103, App
26	37.8	1.0	6840	US-11-090-739-123	Sequence 123, App
C 27	37.6	1.0	600	US-10-750-185-3097	Sequence 3097, App
C 28	37.6	1.0	173602	US-11-121-086-25	Sequence 25, Appli
C 29	37.2	1.0	1821	US-10-750-185-32672	Sequence 32672, A
C 30	37.2	1.0	85682	US-11-117-187-205	Sequence 205, App
C 31	37.2	1.0	611587	US-11-117-187-209	Sequence 209, App
C 32	37	1.0	600	US-10-750-185-1884	Sequence 1884, App
C 33	37	1.0	1213	US-11-112-908-420	Sequence 420, App
C 34	37	1.0	1522	US-10-750-185-44481	Sequence 44481, A
C 35	37	1.0	3313	US-10-750-185-26521	Sequence 26521, A
C 36	37	1.0	31032	US-10-829-826B-23	Sequence 23, Appli
C 37	37	1.0	126552	US-11-121-086-1	Sequence 1, Appli
C 38	37	1.0	164810	US-11-121-086-4	Sequence 4, Appli
C 39	37	1.0	177623	US-11-112-908-41	Sequence 41, Appli
C 40	37	1.0	191684	US-11-121-086-2	Sequence 2, Appli
C 41	36.8	1.0	668	US-10-847-539-2	Sequence 2, Appli
C 42	36.8	1.0	3880	US-10-485-517-9	Sequence 9, Appli
C 43	36.8	1.0	189252	US-11-121-086-54	Sequence 54, Appli
C 44	36.6	1.0	854	US-10-750-185-48779	Sequence 48779, A
C 45	36.6	1.0	1238	US-10-750-185-44723	Sequence 44723, A

ALIGNMENTS

RESULT 1
US-11-121-086-2/c
; Sequence 2, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138, 6000-00000
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: US/11/121, 086
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2
; LENGTH: 191684
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-2

Query Match 1.2%; Score 44.8; DB 7; Length 191684;
Best Local Similarity 60.8%; Pred. No. 1.3;
Matches 73; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 565 AGCCACACACCTTACACATATGCTCGCCGCAAGAGGCCGACCGTCTTCTCC 624
DB 50563 ACCTCACACCGCACGACCTCCATCTCCATCACTCCACCTCCAGCATCTCC 50504
QY 625 GACTTCATTTCACCTCGATTCCTCTCCCTCCCTTCCGACCGACCGACCGCC 684
DB 50503 TCCATCACTTCACCTTCATTCACACACCTCCACCTCCAGCACTCCACCTCCACC 50444

RESULT 2
US-11-121-086-61/c
; Sequence 61, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138, 6000-00000
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: US/11/121, 086
; PRIOR FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570

```
;; PRIOR FILING DATE: 2004-05-04
;; NUMBER OF SEQ ID NOS: 107
;; SOFTWARE: PatentIn version 3.3
;; SEQ ID NO 61
;; LENGTH: 169495
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: modified_base
;; LOCATION: (70072)..(70171)
;; OTHER INFORMATION: a, c, g, t, unknown or other
;; NAME/KEY: modified_base
;; LOCATION: (139457)..(157244)
;; OTHER INFORMATION: a, c, g, t, unknown or other
US-11-121-086-61
```

```
Query Match 1.2%; Score 42.4; DB 7; Length 169495;
Best Local Similarity 48.7%; Pred. No. 5.5;
Matches 115; Conservative 0; Mismatches 121; Indels 0; Gaps 0;
```

```
Qy 2389 GTGTTAGTAGTAGACCTGTTGGTAAATGAGGCGGATGAGCGCTGGTCTTT 2448
Db 67566 GTGTTGCTGCTGATGCTGTTGATATCTGTTGGTGGTCTTTGCTGCTT 67507
Qy 2449 ATTGCAAGCTGTAGAACCCTTGAAACTTGAATGATTAATGCAATTCGAGCT 2508
Db 67506 ACTGTTGGTGTACTCTGTTGTTGACGCTGTTGATGCTGCTGTTGATGAT 67447
Qy 2509 GGGCTCTAGAGAGAGAGGTTGATGAATCTACTGTTGAAGTCCGTTGCTGATATGTA 2568
Db 67446 GCTGTGATGATGCTGTTGATGCTGCTGTTGATGATGCTGTTGATGATG 67387
Qy 2569 AAGAGGCAAGTGTGAAGATCTAGCTGCTGGTGGCAATTTGACATTTTCACT 2624
Db 67386 ATTGCTGCTTTTGGTGGTGTGCTGCTGCTGATGAGTGGTGGTGGTGGTCT 67331
```

```
RESULT 3
US-10-750-185-40551/c
;; Sequence 40551, Application US/10750185
;; Publication No. US20050260603A1
;; GENERAL INFORMATION:
;; APPLICANT: MMI GENOMICS, INC.
;; APPLICANT: DENISE, Sue K.
;; APPLICANT: KERR, Richard
;; APPLICANT: ROSENFELD, David
;; APPLICANT: HOLM, Tom
;; APPLICANT: BATES, Stephen
;; APPLICANT: FANTIN, Dennis
;; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
;; FILE REFERENCE: MM1100-2
;; CURRENT APPLICATION NUMBER: US/10/750,185
;; CURRENT FILING DATE: 2003-12-31
;; PRIOR APPLICATION NUMBER: US 60/437,482
;; PRIOR FILING DATE: 2002-12-31
;; NUMBER OF SEQ ID NOS: 64922
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO 40551
;; LENGTH: 1199
;; TYPE: DNA
;; ORGANISM: Bovine
US-10-750-185-40551
```

```
Query Match 1.1%; Score 42; DB 6; Length 1199;
Best Local Similarity 50.5%; Pred. No. 0.26;
Matches 102; Conservative 0; Mismatches 100; Indels 0; Gaps 0;
```

```
Qy 2679 GATTTCCTTATGATCTGATGCTACATAGATGATTAATGATGCAATTTTCA 2738
Db 2732 GTTCTCTTTTGGAACTGATGCTGCTGATTAATGCTGAGAAATACATTAATGT 213
Qy 2739 TATATCTGATTTCTCAAAATATGCTTTGTTTGAGCTAAGAACATGTTCCACTTA 2798
```

```
Db 212 TTTCCTTTCTTAATTTTCAATTCCTGTTTCTGCTCTCAATTCAGAGATGTTCCGT 153
Qy 2799 TACATGTCCTCAAAAGTTGACCAAGATTAACAGTTGCTGATTAATTTGATTAATG 2858
Db 152 TTCAATGCTTAATTTTATATCTGAGTTATCAATTTTGGTAAATATTTTACTTTTCAA 93
Qy 2859 CTGCTGAATTTTGTGATCAAA 2880
Db 92 GAGTTCTTTTATTCCTTAA 71
```

```
RESULT 4
US-11-121-086-75/c
;; Sequence 75, Application US/11121086
;; Publication No. US20050266459A1
;; GENERAL INFORMATION:
;; APPLICANT: FOLISEN, TIM S.
;; APPLICANT: NIELSEN, KIRSTEN V.
;; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
;; FILE REFERENCE: 09138.6000-00000
;; CURRENT APPLICATION NUMBER: US/11/121,086
;; CURRENT FILING DATE: 2005-05-04
;; PRIOR APPLICATION NUMBER: 60/567,570
;; PRIOR FILING DATE: 2004-05-04
;; NUMBER OF SEQ ID NOS: 107
;; SOFTWARE: PatentIn version 3.3
;; SEQ ID NO 75
;; LENGTH: 161874
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-11-121-086-75
```

```
Query Match 1.1%; Score 40.6; DB 7; Length 161874;
Best Local Similarity 47.8%; Pred. No. 17;
Matches 116; Conservative 0; Mismatches 129; Indels 0; Gaps 0;
```

```
Qy 2753 TCAAAATATGCTGTTTGTGAGCTAAGAACATAGTTCCACTTAATACATGTCCTCAAA 2812
Db 49935 TCATTAATGATTTATTAATCTAAGATTAATTAATTAATTAATTAATTAATTAATTA 49876
Qy 2813 GTTGTACCAAGATTAACAAATTTGCTGATTAATTTTCACTAATTAATGCTGTAATTTT 2872
Db 49875 AATGTTTCAATGCTCACTGCTGACAGCAAGATTAATTAATTAATTAATTAATTAAT 49816
Qy 2873 TGATCAAACTGTGACAGCAAGATTAATTAATTTCACTCAATTTCTGTTAGATTAAGT 2932
Db 49815 TGATGAATTAATGATGATTTAATTTTAACAGACATGCAAACTTTTCTGATTAATGATCA 49756
Qy 2933 AGGATTAAGATTTGCTTAATGCTGCTTTGCTCAACTTTTCTTCTGATTTTCTT 2992
Db 49755 TATGTAATCATTTTCACTTAATGATGATCACTCAACTTAATTTGATGACTACTATCTC 49696
Qy 2993 TTGCATT 2999
Db 49695 TGCCATT 49689
```

```
RESULT 5
US-10-829-826B-21
;; Sequence 21, Application US/10829826B
;; Publication No. US20050266397A1
;; GENERAL INFORMATION:
;; APPLICANT: Ecker, David J.
;; APPLICANT: Hostadler, Steven A.
;; APPLICANT: Sampath, Rangarajan
;; APPLICANT: Blyn, Lawrence B.
;; APPLICANT: Hall, Thomas A.
;; APPLICANT: Massire, Christian
;; TITLE OF INVENTION: METHOD FOR IDENTIFICATION OF CORONAVIRUSES
;; FILE REFERENCE: IBIS0075-100 (IBIS-0058US)
;; CURRENT APPLICATION NUMBER: US/10/829,826B
;; CURRENT FILING DATE: 2004-04-22
```


Db 3700 GATGCTATGTTCTTTTATGAGTGTGTCACATGTCGAACTGTGAGTCTATG 3759
Qy 2443 GTCTTATTTGAGAGAGTGTAGACCTCTGTGAAACTTTGAATATATGCAAT 2502
Db 3760 GTACTTATTTGATGTTGATGTGTCATTTACAGCCCACTTGTCTTTAAGATATAGTTGTT 3819
Qy 2503 CGAGCTGGGGTCTCAGAGATGACGTTGATGAAACTACTGTTGAAGTCCGTTGCTGAT 2562
Db 3820 TGTGATTTATTAACGCTAGTGTGTATTAAGACGCTGTGTGTGATGTTAATGAT 3879
Qy 2563 A 2563
Db 3880 A 3880

RESULT 9

US-11-140-417-3
; Sequence 3, Application US/11140417
; Publication No. US20050266515A1
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Deborah A
; APPLICANT: Eddy, Edward M
; TITLE OF INVENTION: GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE-S (GAPDS), A GLYCOL TIC
; TITLE OF INVENTION: ENZYME EXPRESSED ONLY IN MALE GERM CELLS, IS A TARGET FOR M LE
; FILE REFERENCE: 421/76/2 PCT/CIP
; CURRENT FILING DATE: 2005-05-27
; PRIOR FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: US 60/429,638
; PRIOR FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: PCT/US2003/037800
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3
; LENGTH: 1450
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (4)..(1302)
US-11-140-417-3

Query Match 1.1%; Score 39.8; DB 7; Length 1450;
Best Local Similarity 51.4%; Pred. No. 1.2;
Matches 92; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

Qy 507 TGGTCTCTCCCATTCCTCAATTAATGCGGATTAACACCGGCGAGCAAGCTCCGACCTAG 566
Db 107 TGATTCAGCACCTTCACCTCCACCCAGTTGAGAGCTTCGCGACCAAGAGAGAGC 166
Qy 567 CCACAAACACTCTCAACTATCTGCTCCGACGAAATGAGCCGACGCTTCTCTCCGA 626
Db 167 CGCGCCACCGCCGACCAACCGCTCTCTCAACCCAGATGAGCCAGAGAGAGCTTAAAGG 226
Qy 627 CTTCATTTACCTCTCCGATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 685
Db 227 CTCACCTCCACCTTCACCT 285

RESULT 10

US-11-121-086-3/c
; Sequence 3, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138,6000-0000
; CURRENT FILING DATE: 2005-05-04
; PRIOR FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04

; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 3
; LENGTH: 168516
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-3

Query Match 1.1%; Score 39.8; DB 7; Length 168516;
Best Local Similarity 56.5%; Pred. No. 28;
Matches 74; Conservative 0; Mismatches 57; Indels 0; Gaps 0;

Qy 556 CTCGACGTAGCCACCAACCTCTACACTATCTGCTCCGCGCAAGAGCCGACCGT 615
Db 68578 CTCCTACTCTCTCCCTTGGCTCCCTCCCTCCCTCTTACTCTCTCCCTCTCTCTCTCT 68519
Qy 616 CTCTCTCGACCTTCATTTACCTCGATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 675
Db 68518 TCTCTCTCGGCTTCACCTCTCTGCGCTTCCCTCCCTCTCTCTCTCTCTCTCTCTCTCTCT 68459

Qy 676 ACCACCGCCAC 686
Db 68458 CCTCTCTCTC 68448

RESULT 11

US-10-750-185-3575
; Sequence 3575, Application US/10750185
; Publication No. US2005026603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM11100-2
; CURRENT FILING DATE: 2003-12-31
; PRIOR FILING DATE: 2002-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3575
; LENGTH: 600
; TYPE: DNA
; ORGANISM: Bovine
US-10-750-185-3575

Query Match 1.1%; Score 39.6; DB 6; Length 600;
Best Local Similarity 50.0%; Pred. No. 0.74;
Matches 99; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

Qy 289 ATTTAATCATATAAGTCGTCGACGTGATACATCATAGATTGTTTTTATTTTTC 348
Db 4 ATATCACTTATATCAACCCCTGAATACTAGCTAGCAAACTGTGAGACTATTAAT 63
Qy 349 GTAGCTGTATGTTTTTTTATTTAATTACTATCTCAAAATCAAAATTCATTAACC 408
Db 64 ATGATGTATATATATATACAGGTGAAAGTTTACGAGAAATAATTTTTCATATGAAA 123
Qy 409 TAGACGACCAAGCTCTTCAATATGTAAACAGAACTTTTGTAGAGCTTA 468
Db 124 AATTAAGCAATATATATTTTATTTATCTTAAGATATATAGCATGTGACTTATAT 183
Qy 469 AAAGACCTCCCATGAA 486
Db 184 AAAAGATTTCTGTGAAA 201

RESULT 12

; TYPE: DNA
; ORGANISM: Mus musculus
US-11-140-417-22

Query Match 1.1%; Score 39.2; DB 7; Length 11462;
Best Local Similarity 51.7%; Pred. No. 6.8;
Matches 89; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

OY	514	TCGCCATTCCCAATTATGCGGATTTACCA	CGGCGAGACAAAGCTCCGAGTACCCACAAC	573
DB	3221	TCACACACCCACAGGTTGAAGAAGACAGCA	CGCCACCGCCGCCACTCCACTCCACCC	3280
OY	574	ACCTTACAACTATCTGCTCCGCCAGCAAA	TGGCCGACCGGCTTCTCTCCGACTTCAAT	633
DB	3281	ACCACTCTCTCTCTCTCCGCCAGATAGAG	CCAGACAAGTTGAAGAAGGCTCCCCC	3340
OY	634	TTCACTTCGATTCCTCTCTCTCTCTCTCT	CGCCACCGCCACCAACCGCA	685
DB	3341	TCCTCTCTCTCTCTCTCTCTCTCTCTCT	CTCTCTCTCTCTCTCTCTCTCTCTCTCT	3392

Search completed: December 10, 2005, 18:24:17
Job time : 325.571 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 10, 2005, 06:08:12 ; Search time 2797.5 Seconds
(without alignments)
10839.614 Million cell updates/sec

Title: US-10-600-070b-3
Perfect score: 3667
Sequence: 1 tgcctgcattcaaggaagaat.....ctatacataagggctacaa 3667

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications_NA_Main:*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3667	100.0	3667	US-10-600-070-3	Sequence 3, Appli
2	3663.8	99.9	3667	US-10-600-070-10	Sequence 10, Appl
3	1968.4	53.7	2679	US-10-739-930-227	Sequence 227, App
4	1931	52.7	2637	US-10-600-070-130	Sequence 130, App
5	1724	47.0	2406	US-10-600-070-1	Sequence 1, Appli
6	1720.8	46.9	2406	US-10-600-070-9	Sequence 9, Appli
7	1717.6	46.8	2406	US-10-600-070-128	Sequence 128, App
8	402.2	11.0	561	US-10-600-070-132	Sequence 132, App
9	295.6	8.1	631	US-10-600-070-184	Sequence 184, App
10	262.4	7.2	1146	US-10-424-599-119007	Sequence 125007,
11	208.8	5.7	660	US-10-600-070-135	Sequence 135, App
12	194	5.3	2283	US-10-600-070-126	Sequence 126, App
13	178.6	4.9	607	US-10-021-323-3562	Sequence 3562, App
14	175	4.8	439	US-09-732-627A-2154	Sequence 2154, App
15	161	4.4	545	US-10-600-070-175	Sequence 175, App
16	159.4	4.3	552	US-10-021-323-3536	Sequence 3536, App
17	158	4.3	608	US-10-600-070-137	Sequence 137, App
18	149.4	4.1	1411	US-10-425-115-81853	Sequence 81853, A
19	146.2	4.0	2130	US-10-437-963-69932	Sequence 69932, A
20	145.6	4.0	527	US-10-600-070-134	Sequence 134, App
21	145.4	4.0	647	US-10-767-701-4069	Sequence 4069, App
22	141	3.8	1039	US-10-424-599-35059	Sequence 35059, A
23	139.8	3.8	1536	US-10-425-115-57452	Sequence 57452, A

24	136.2	3.7	647	7	US-10-600-070-185	Sequence 185, App
25	134.6	3.7	652	7	US-10-600-070-186	Sequence 186, App
26	133.2	3.6	553	7	US-10-600-070-178	Sequence 178, App
27	130.2	3.6	446	7	US-10-600-070-151	Sequence 151, App
28	127.6	3.5	479	7	US-10-600-070-150	Sequence 150, App
29	125	3.4	537	7	US-10-600-070-143	Sequence 143, App
30	122.6	3.3	307	7	US-10-600-070-138	Sequence 138, App
31	118.6	3.2	871	7	US-10-600-070-153	Sequence 153, App
32	117.8	3.2	1703	7	US-10-437-963-69933	Sequence 69933, A
33	110.4	3.0	618	7	US-10-437-963-8187	Sequence 8187, A
34	93.6	2.6	491	7	US-10-600-070-174	Sequence 174, App
35	92.6	2.5	187	7	US-10-600-070-136	Sequence 136, App
36	92	2.5	418	7	US-10-600-070-144	Sequence 144, App
37	91.4	2.5	460	7	US-10-600-070-187	Sequence 187, App
38	90.4	2.5	420	7	US-10-600-070-176	Sequence 176, App
39	89.4	2.4	604	7	US-10-600-070-147	Sequence 147, App
40	89.4	2.4	653	7	US-10-600-070-148	Sequence 148, App
41	89.4	2.4	1032	7	US-10-767-701-9474	Sequence 9474, App
42	87.2	2.4	480	7	US-10-600-070-145	Sequence 145, App
43	86.8	2.4	622	7	US-10-600-070-146	Sequence 146, App
44	85.2	2.3	300	7	US-10-600-070-180	Sequence 180, App
45	83	2.3	541	7	US-10-600-070-154	Sequence 154, App

ALIGNMENTS

RESULT 1
US-10-600-070-3
; Sequence 3, Application US/10600070
; Publication No. US20040139500A1
; GENERAL INFORMATION:
; APPLICANT: Osteoryoung, Katherine W.
; APPLICANT: Vichra, Stanislaw
; APPLICANT: Koksharova, Olga A.
; APPLICANT: Geo, Hongo
; TITLE OF INVENTION: Placid Division and Related Genes and Proteins, and Methods of
; TITLE OF INVENTION: Use
; FILE REFERENCE: MSU-08153
; CURRENT APPLICATION NUMBER: US/10/600,070
; CURRENT FILING DATE: 2003-06-20
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 3667
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; US-10-600-070-3

Query Match 100.0%; Score 3667; DB 7; Length 3667;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3667; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TGTTCGATTAGAGAAATGCAATTTAGCAATTTGCTGATTTTCACAGATTTTG 60
DB 1 TGTTCGATTAGAGAAATGCAATTTAGCAATTTGCTGATTTTCACAGATTTTG 60
QY 61 CTGGCTATAGGATTCATGCTGCTGCTTTTCAATTTACATGTCATATAGTTTCG 120
DB 61 CTGGCTATAGGATTCATGCTGCTGCTTTTCAATTTACATGTCATATAGTTTCG 120
QY 121 AATTTACACATTTTCAGTTGATTTTAAGAAAGAGAGGAAATGATGGGTTTGTGGG 180
DB 121 AATTTACACATTTTCAGTTGATTTTAAGAAAGAGAGGAAATGATGGGTTTGTGGG 180
QY 181 TTTTAACCTTTAAATAGTCAAGAAATTAAGTCAATTTGCTGCTATATAGTTTAA 240
DB 181 TTTTAACCTTTAAATAGTCAAGAAATTAAGTCAATTTGCTGCTATATAGTTTAA 240
QY 241 AATGAGGCACTCCACAGGTTCTTAGTGGAATAGATTATTTAGAGATTTTACATCAT 300
DB 241 AATGAGGCACTCCACAGGTTCTTAGTGGAATAGATTATTTAGAGATTTTACATCAT 300

Qy 301 AAAGTCCGTGGCGAGTGTAACTCATTAAGTTGTTTTTATTTTTCAGTACGTGGTAT 360
Db 301 AAAGTCCGTGGCGAGTGTAACTCATTAAGTTGTTTTTATTTTTCAGTACGTGGTAT 360
Qy 361 GTTTTTGAATTTAACTTAATCTACTCAAAATCAAAATTCATTAACCCCTAGACGCAAA 420
Db 361 GTTTTTGAATTTAACTTAATCTACTCAAAATTCATTAACCCCTAGACGCAAA 420
Qy 421 CAGTCTCTTCAATATGTAAAAACAGAACAAAGTTTTTGTATAGTACCTTAAAAAGACATCCC 480
Db 421 CAGTCTCTTCAATATGTAAAAACAGAACAAAGTTTTTGTATAGTACCTTAAAAAGACATCCC 480
Qy 481 ATGGAAGCTGTAGAGTACGTGGGCACTTGGTCTCTCCCATTCCTCAATTAATGCCATTACCA 540
Db 481 ATGGAAGCTGTAGAGTACGTGGGCACTTGGTCTCTCCCATTCCTCAATTAATGCCATTACCA 540
Qy 541 CCGGCGAGCAAAAGCTCGAGAGTAGCAACAACCTCTACAACTATCTGCTCGGCGAGC 600
Db 541 CCGGCGAGCAAAAGCTCGAGAGTAGCAACAACCTCTACAACTATCTGCTCGGCGAGC 600
Qy 601 AAATGGGCGGACCGTCTTCTCCGACTTCAATTTCACTCGGATTCCTCTCTCTCTCC 660
Db 601 AAATGGGCGGACCGTCTTCTCCGACTTCAATTTCACTCGGATTCCTCTCTCTCTCC 660
Qy 661 TTCCGCAACCGGCAACCAACGCGGCACTCTCGTCTCTGCGCAACATCTATGTATGCTCC 720
Db 661 TTCCGCAACCGGCAACCAACGCGGCACTCTCGTCTCTGCGCAACATCTATGTATGCTCC 720
Qy 721 GAAGCGCAAGTCCCATCCCATTTGATTTCTACAGAGTATTAGAGGCTCAAAACATTTTC 780
Db 721 GAAGCGCAAGTCCCATCCCATTTGATTTCTACAGAGTATTAGAGGCTCAAAACATTTTC 780
Qy 781 TTAAACGAGTGAATCAGAAAGAGCATTCGAGCTAGGGTTTGAAACCGGCGCAATTCGGT 840
Db 781 TTAAACGAGTGAATCAGAAAGAGCATTCGAGCTAGGGTTTGAAACCGGCGCAATTCGGT 840
Qy 841 TTCAGCGAGAGCGCTTTATATCAGCCGAGAGAGATTTCAAGCTCTTGAGGAAATCTGTG 900
Db 841 TTCAGCGAGAGCGCTTTATATCAGCCGAGAGAGATTTCAAGCTCTTGAGGAAATCTGTG 900
Qy 901 TCTAATCTCTGGTCTAGAGAGAGTACAAATGAAAGGCTTCTTGATGATGAAGAGCTACA 960
Db 901 TCTAATCTCTGGTCTAGAGAGAGTACAAATGAAAGGCTTCTTGATGATGAAGAGAGCTACA 960
Qy 961 GTCATCACTGATGTTCTTGGGATTAAGGTAAATTCGATTTGGAATTAAGATTTCTTC 1020
Db 961 GTCATCACTGATGTTCTTGGGATTAAGGTAAATTCGATTTGGAATTAAGATTTCTTC 1020
Qy 1021 GTTTTAAATTTGATGAATTTGATTAAGAGAAAGAACTTTTATCTAGTGAAGGTTCTCGGGC 1080
Db 1021 GTTTTAAATTTGATGAATTTGATTAAGAGAAAGAACTTTTATCTAGTGAAGGTTCTCGGGC 1080
Qy 1081 TCTCTGTATTTGCAAGAGGTGTGAGACTGAGATAGTTCCTCGGGTGTGGAGGCTCT 1140
Db 1081 TCTCTGTATTTGCAAGAGGTGTGAGACTGAGATAGTTCCTCGGGTGTGGAGGCTCT 1140
Qy 1141 GCTTAAAGAGAGGTTGCTTAAGTGTGTTTAAAGCAAGATGTGTTTATGATTAAGCGCTTC 1200
Db 1141 GCTTAAAGAGAGGTTGCTTAAGTGTGTTTAAAGCAAGATGTGTTTATGATTAAGCGCTTC 1200
Qy 1201 GTTTCCTCATGTCTGAGAGGATGCTATAGGCAATGGAATCACTGATTTTATTAATCTGGTTA 1260
Db 1201 GTTTCCTCATGTCTGAGAGGATGCTATAGGCAATGGAATCACTGATTTTATTAATCTGGTTA 1260
Qy 1261 TGAAGTTGTGAGGAAGCTTTGAAGCTTTTACAGGTATGTTGACTTGCTTTGGTAAATTTG 1320
Db 1261 TGAAGTTGTGAGGAAGCTTTGAAGCTTTTACAGGTATGTTGACTTGCTTTGGTAAATTTG 1320
Qy 1321 ACGAGCGTTGCTTTAATAAGAACTTTCTTGATTTGATACCTTTGATTAATGAGCTTTGTGTA 1380
Db 1321 ACGAGCGTTGCTTTAATAAGAACTTTCTTGATTTGATACCTTTGATTAATGAGCTTTGTGTA 1380
Qy 1381 GGAGGAAGGAGCAAGTACCTTGCACCGGATTTTACGTGCACAAATGTATGATGACTTTGGA 1440

Db 1381 GGAGGAAGGAGCAAGTACCTTGCACCGGATTTTACGTGCACAAATGTATGATGACTTTGGA 1440
Qy 1441 AGAGTACATCCCGGTTATGCTTGGAGTACTTGGCTTACCGCTGTGATGATTAACGC 1500
Db 1441 AGAGTACATCCCGGTTATGCTTGGAGTACTTGGCTTACCGCTGTGATGATTAACGC 1500
Qy 1501 TGGAAAAAGCTAAATGTTTTAAGCGGTGTGCGGAATATTTTGTGTCTGTGTGAGAGAG 1560
Db 1501 TGGAAAAAGCTAAATGTTTTAAGCGGTGTGCGGAATATTTTGTGTGTGTGAGAGAG 1560
Qy 1561 TGGAGCACTCAGCTCTTGTGGGGGTTTGAACCGGAGAAATTTATATGATAGAGGTTTTT 1620
Db 1561 TGGAGCACTCAGCTCTTGTGGGGGTTTGAACCGGAGAAATTTATGATAGAGGTTTTT 1620
Qy 1621 ACGAATACAGCTGCTGAGAGGATACAGTTTATAGATTAATCTTTTAAATTTCTTTAGCA 1680
Db 1621 ACGAATACAGCTGCTGAGAGGATACAGTTTATAGATTAATCTTTTAAATTTCTTTAGCA 1680
Qy 1681 TGATATACTTTAGGTTTCTCATTTTATATGATATGTGTGTAGAGTTGATCTTTTGTGA 1740
Db 1681 TGATATACTTTAGGTTTCTCATTTTATATGATATGTGTGTAGAGTTGATCTTTTGTGA 1740
Qy 1741 GCTAACCCCAAGCAATATTTCCAGAGAGTCAATTTGAAGTTTACGAAGTTGCACTTGCCTT 1800
Db 1741 GCTAACCCCAAGCAATATTTCCAGAGAGTCAATTTGAAGTTTACGAAGTTGCACTTGCCTT 1800
Qy 1801 GTGGCTCAAGCTTTATTTGTTAAGAAAGCACACTTTTACAGAGTGTGATTAAGCAATTC 1860
Db 1801 GTGGCTCAAGCTTTATTTGTTAAGAAAGCACACTTTTACAGAGTGTGATTAAGCAATTC 1860
Qy 1861 CAGCACTTACAGAGGTTAAGTATGCTATGAGATTCCTGCGATGTTGTATGATACA 1920
Db 1861 CAGCACTTACAGAGGTTAAGTATGCTATGAGATTCCTGCGATGTTGTATGATACA 1920
Qy 1921 CGGAATATTTGGAGATGAGCTTCCGTCTAAGAAAGGGAATCTGTGACGCTTATAGGC 1980
Db 1921 CGGAATATTTGGAGATGAGCTTCCGTCTAAGAAAGGGAATCTGTGACGCTTATAGGC 1980
Qy 1981 AAAGTTGATGATATCCCTATATGTTGGGCTTTAGACAGTGAAGATTCACAAATATAGAAAT 2040
Db 1981 AAAGTTGATGATATCCCTATATGTTGGGCTTTAGACAGTGAAGATTCACAAATATAGAAAT 2040
Qy 2041 CCAAGCTATTTGAGATTTGTTTGGAGAAATTCAAATGTATGACAAATGATCTCCCT 2100
Db 2041 CCAAGCTATTTGAGATTTGTTTGGAGAAATTCAAATGTATGACAAATGATCTCCCT 2100
Qy 2101 GGAATATGCAAAATTTGTGAAACCTGTGTGGCAGGGGTGTCTTCTAGGTTCAAGAGAC 2160
Db 2101 GGAATATGCAAAATTTGTGAAACCTGTGTGGCAGGGGTGTCTTCTAGGTTCAAGAGAC 2160
Qy 2161 ACCAAGATTAATAATTTAAATCTGGGGACTAATGATGATCTATGTTTTGATTAAC 2220
Db 2161 ACCAAGATTAATAATTTAAATCTGGGGACTAATGATGATCTATGTTTTGATTAAC 2220
Qy 2221 TTGGAAGAAGTGAAGGATGTTCAAGGTTCTCTTTAGCTGCTGCACTATAGGCAAG 2280
Db 2221 TTGGAAGAAGTGAAGGATGTTCAAGGTTCTCTTTAGCTGCTGCACTATAGGCAAG 2280
Qy 2281 ATTGAGCCGAGATGTAAGAAAGCTATAGTCTATGCAAGCACTGAGAAAGTTTTCTTC 2340
Db 2281 ATTGAGCCGAGATGTAAGAAAGCTATAGTCTATGCAAGCACTGAGAAAGTTTTCTTC 2340
Qy 2341 CGCTATACAGATGAATACTCGGCTGAACCCAGAGATGTGCAAGAGACAGTGTTAAGTGA 2400
Db 2341 CGCTATACAGATGAATACTCGGCTGAACCCAGAGATGTGCAAGAGACAGTGTTAAGTGA 2400
Qy 2401 GATCTGTGTGTAACTATGTAAGCCGTGATGTGAGCTGTGTCTTTATTTGCAAGACT 2460
Db 2401 GATCTGTGTGTAACTATGTAAGCCGTGATGTGAGCTGTGTCTTTATTTGCAAGACT 2460
Qy 2461 GTAAGACCTCTGAAAACTTTGAACCTAATGATTAAGCAATTTGAGAGTGGGGTCTCGAG 2520
Db 2461 GTAAGACCTCTGAAAACTTTGAACCTAATGATTAAGCAATTTGAGAGTGGGGTCTCGAG 2520

Db	2461	GTAAAGACCCCTCTGAAAACTTTGAAACTAATGATTATGCAATTCGAGCTGGGGTCTCAGAG	2520
Qy	2511	AGTAGCCGTTGATGAAACTCACTGTTGAAATGTCGCCGTGCTGATATGTTAAAGAGCAAGT	2580
Db	2521	AGTAGCCGTTGATGAAACTCACTGTTGAAATGTCGCCGTGCTGATATGTTAAAGAGCAAGT	2580
Qy	2581	GTAAGATCTCAGCTGCTGTGTGGCAATTTGAGACTGATTTCACTGTTCAAGCCAGAGAT	2640
Db	2581	GTAAGATCTCAGCTGCTGTGTGGCAATTTGAGACTGATTTCACTGTTCAAGCCAGAGAT	2640
Qy	2641	TTTTCTTAAAGCAGCTCACTCTTTTCMAAGCAAGATATGTTTTCTTATGGAATCTGAT	2700
Db	2641	TTTTCTTAAAGCAGCTCACTCTTTTCMAAGCAAGATATGTTTTCTTATGGAATCTGAT	2700
Qy	2701	GTCGCTACCAATAGGTATGATTTAATGATGCAATTTCAATATCTGCATTCCTCAATA	2760
Db	2701	GTCGCTACCAATAGGTATGATTTAATGATGCAATTTCAATATCTGCATTCCTCAATA	2760
Qy	2761	TGCTTGTTTTGAGCTAAGAACATAGTTCACCTTAATATGTCACCAAAAGTTGATACC	2820
Db	2761	TGCTTGTTTTGAGCTAAGAACATAGTTCACCTTAATATGTCACCAAAAGTTGATACC	2820
Qy	2821	AAGATTTAAACAAGTTGCTGAGTAAATTTCACTAAATATGCTCTGAAATTTTTGATCAAA	2880
Db	2821	AAGATTTAAACAAGTTGCTGAGTAAATTTCACTAAATATGCTCTGAAATTTTTGATCAAA	2880
Qy	2881	CTGTAGACAGAAATGTAAATTTCACTCTCAACATTTCTGTTAGAAATACGTAGATTAAG	2940
Db	2881	CTGTAGACAGAAATGTAAATTTCACTCTCAACATTTCTGTTAGAAATACGTAGATTAAG	2940
Qy	2941	AGATTTGCCTTTGTGTGGCTTTGTCCAACTTTTCTTCTGATATTTTTCTTTTCGATTT	3000
Db	2941	AGATTTGCCTTTGTGTGGCTTTGTCCAACTTTTCTTCTGATATTTTTCTTTTCGATTT	3000
Qy	3001	AGGGTCAGTCAGAGCTGACGATTTAGAAAGCACTTCCAGAAATGAGTCTAGAGACTGAGA	3060
Db	3001	AGGGTCAGTCAGAGCTGACGATTTAGAAAGCACTTCCAGAAATGAGTCTAGAGACTGAGA	3060
Qy	3061	GAATATAGTATCCAAATGAGCAGAAAGATTAAAGTCTGTGGCTTTTGGGCTGATCAACGCGAT	3120
Db	3061	GAATATAGTATCCAAATGAGCAGAAAGATTAAAGTCTGTGGCTTTTGGGCTGATCAACGCGAT	3120
Qy	3121	AGAAATGTTACCAAGGTGAGAGGAAATTAATCTCAATTCATTCATTTGTGTGAAAACTGT	3180
Db	3121	AGAAATGTTACCAAGGTGAGAGGAAATTAATCTCAATTCATTCATTTGTGTGAAAACTGT	3180
Qy	3181	TGGAATGATTAATAGTCTGTGGCTCTGTTGATTTCTGTTATTTAATAGGTTTTGGAATGGGC	3240
Db	3181	TGGAATGATTAATAGTCTGTGGCTCTGTTGATTTCTGTTATTTAATAGGTTTTGGAATGGGC	3240
Qy	3241	GAATGCTGAAGATTGGAATCTGACAGAGCAAGCTGAACTGGCGACAGCTGGGTGTTATG	3300
Db	3241	GAATGCTGAAGATTGGAATCTGACAGAGCAAGCTGAACTGGCGACAGCTGGGTGTTATG	3300
Qy	3301	ATTATTAACACTGTGAAACTATCTGTGACAGTGTGACAGCTCAGCAGATGAAACCCGTG	3360
Db	3301	ATTATTAACACTGTGAAACTATCTGTGACAGTGTGACAGCTCAGCAGATGAAACCCGTG	3360
Qy	3361	CTCAGGTGGAAGCAACTCTGAGAGAGCTGTGTTCTCATCTGATTTGGTTCAATCCAGAAA	3420
Db	3361	CTCAGGTGGAAGCAACTCTGAGAGAGCTGTGTTCTCATCTGATTTGGTTCAATCCAGAAA	3420
Qy	3421	ACAATGCTACAGATGTGGAACCTTACACAACAATATGCAAGATTTTCTGGTCCAGTCAAG	3480
Db	3421	ACAATGCTACAGATGTGGAACCTTACACAACAATATGCAAGATTTTCTGGTCCAGTCAAG	3480
Qy	3481	GGTGGAAATATCATGGAAGGCTCTGTCTTGATCATATAATATCTCATATGATGACATGTCT	3540
Db	3481	GGTGGAAATATCATGGAAGGCTCTGTCTTGATCATATAATATCTCATATGATGACATGTCT	3540
Qy	3541	GAGCTTGAGAGATTTCTGTTGTTCTGTGAAATTTCTCTCTAAGTAAATGTTTAAATATGA	3600
Db	3541	GAGCTTGAGAGATTTCTGTTGTTCTGTGAAATTTCTCTCTAAGTAAATGTTTAAATATGA	3600

OY		3601	ACACAAAAAATTAAAGTCTTGCGCACACCCTTTCTTCATCTAAACTTAACATAAGG	3660
Dd		3601	ACACAAAAAATTAAAGTCTTGCGCACACCCTTTCTTCATCTAAACTTAACATAAGG	3660
OY		3661	GCTACAA 3667	
Dd		3661	GCTACAA 3667	
 RESULT 2 US-10-600-070-10				
; Sequence 10, Application US/10600070				
; Publication No. US20040139500A1				
; GENERAL INFORMATION:				
; APPLICANT: Oesteryoung, Katherine W.				
; APPLICANT: Vitna, Stanislav				
; APPLICANT: Koksharova, Olga A.				
; APPLICANT: Gao, Hongbo				
; TITLE OF INVENTION: Plastid Division and Related Genes and Proteins, and Methods of				
; FILE REFERENCE: MSU-08153				
; CURRENT APPLICATION NUMBER: US/10/600,070				
; CURRENT FILING DATE: 2003-06-20				
; NUMBER OF SEQ ID NOS: 206				
; SOFTWARE: PatentIn version 3.2				
; SEQ ID NO 10				
; LENGTH: 3667				
; TYPE: DNA				
; ORGANISM: Arabidopsis thaliana				
US-10-600-070-10				
 Query Match 99.9%; Score 3663.8; DB 7; Length 3667;				
Best Local Similarity 99.9%; Pred. No. 0;				
Matches 3665; Conservative 0; Mismatches 2; Indels 0; Gaps 0;				
OY		1	TGTTCTGCATTAAAGAGAATACAATTATTAAGCAATTGTCTTGATTTCAACAAGATTTTG	60
Dd		1	TGTTCTGCATTAAAGAGAATACAATTATTAAGCAATTGTCTTGATTTCAACAAGATTTTG	60
OY		61	CTTGCGCATAGGAATCATTGGCTCTGTTGCTTTTACATTTACATGTCATATNAGTTTCG	120
Dd		61	CTTGCGCATAGGAATCATTGGCTCTGTTGCTTTTACATTTACATGTCATATNAGTTTCG	120
OY		121	AATTTTACACATTTAGTTCAGTTGAGTGTAAAGAAAAGAGAGGGAATTGATGGGGTTTTGGGG	180
Dd		121	AATTTTACACATTTAGTTCAGTTGAGTGTAAAGAAAAGAGAGGGAATTGATGGGGTTTTGGGG	180
OY		181	TTTAAACTTTAAAGTAGTCACAGAAATTAAGTCATTGGTTTACTGTCTCTATATGTGTAA	240
Dd		181	TTTAAACTTTAAAGTAGTCACAGAAATTAAGTCATTGGTTTACTGTCTCTATATGTGTAA	240
OY		241	AATGAGGCCAACCTCCAGCGTTCTTAGTGGTAATAGATTATTAGACGATTTAAACATCAT	300
Dd		241	AATGAGGCCAACCTCCAGCGTTCTTAGTGGTAATAGATTATTAGACGATTTAAACATCAT	300
OY		301	AAAGCGGTGGGCACTGTATACATCAPAGATTGTTTTTATTTTTTTCAGTAGCGGTGAT	360
Dd		301	AAAGCGGTGGGCACTGTATACATCAPAGATTGTTTTTATTTTTTTCAGTAGCGGTGAT	360
OY		361	GTTTTTGATTTTAACTTATTAAGTCTCAAATAATCAAAATTCATTAACCTTAGAGACCCAA	420
Dd		361	GTTTTTGATTTTAACTTATTAAGTCTCAAATAATCAAAATTCATTAACCTTAGAGACCCAA	420
OY		421	CAGTCTCTTCAATATGTATTAACAGAACAAAGTTTTTGTATGTAGCTTAAAAAGACCTCCC	480
Dd		421	CAGTCTCTTCAATATGTATTAACAGAACAAAGTTTTTGTATGTAGCTTAAAAAGACCTCCC	480
OY		481	ATGGAAGCTCTGAGTCACTGCGGCAATGGTCTCTCCCCATTCCTCAATTAATGCGCATTAACA	540
Dd		481	ATGGAAGCTCTGAGTCACTGCGGCAATGGTCTCTCCCCATTCCTCAATTAATGCGCATTAACA	540
OY		541	CCGGGACGACAAAGCTCCGAGTACGCCAACAACCTTACAACATATCTGCTCCGCGACG	600

```
|||||
Db 541 CCGGCGACGACAAAGCTCCGACGTAGCCACAACCTCTACACATCTGCTCCGCGACG 600
Qy AAATGGGCGGACCGCTCTCTCCGACCTTCAATTTACCTCGATTCCTCCGCTCC 660
Db 601 AAATGGGCGGACCGCTCTCTCCGACCTTCAATTTACCTCGATTCCTCCGCTCC 660
Qy 601 AAATGGGCGGACCGCTCTCTCCGACCTTCAATTTACCTCGATTCCTCCGCTCC 660
Db 661 TTGCGCACCGGACCGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCG 720
Qy 661 TTGCGCACCGGACCGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCG 720
Db 661 TTGCGCACCGGACCGACCGGACCGGACCGGACCGGACCGGACCGGACCGGACCG 720
Qy 721 GAAGCGCACGCTCCCATCTCCCATCTGATTTCTACAGGATTTAGAGAGCTCAACACATTC 780
Db 721 GAAGCGCACGCTCCCATCTCCCATCTGATTTCTACAGGATTTAGAGAGCTCAACACATTC 780
Qy 721 GAAGCGCACGCTCCCATCTCCCATCTGATTTCTACAGGATTTAGAGAGCTCAACACATTC 780
Db 721 GAAGCGCACGCTCCCATCTCCCATCTGATTTCTACAGGATTTAGAGAGCTCAACACATTC 780
Qy 781 TTAACCGATGGAATCAGAAAGACATTCGAGCTAGGGTTTGAAAACCGCGCAATTCGGT 840
Db 781 TTAACCGATGGAATCAGAAAGACATTCGAGCTAGGGTTTGAAAACCGCGCAATTCGGT 840
Qy 841 TTCAGCGACGACGCTTTTATCATGACCGGACGAGATTTCTCAAGCTGCTTGCGAAACTCTG 900
Db 841 TTCAGCGACGACGCTTTTATCATGACCGGACGAGATTTCTCAAGCTGCTTGCGAAACTCTG 900
Qy 901 TCTAATCTCGGCTCTAGAGAGAGATCATGAGGTCTTCTTGATGATGAAGAGCTACA 960
Db 901 TCTAATCTCGGCTCTAGAGAGAGATCATGAGGTCTTCTTGATGATGAAGAGCTACA 960
Qy 961 GTCACTACATGATGTTCTTGGAATPAAGGTAATTTGATTCGGAATTAAGTTTCTTC 1020
Db 961 GTCACTACATGATGTTCTTGGAATPAAGGTAATTTGATTCGGAATTAAGTTTCTTC 1020
Qy 1021 GTTTTAAATTTTATGATGATTAAGAGAACTTTTATCTAGTGAAGTTCTCTGCGGC 1080
Db 1021 GTTTTAAATTTTATGATGATTAAGAGAACTTTTATCTAGTGAAGTTCTCTGCGGC 1080
Qy 1081 TCTCTGTATATTTGCAAGAGAGTGTGAGACTGAGATAGTTCTTGCGGTTGGTGGCTCT 1140
Db 1081 TCTCTGTATATTTGCAAGAGAGTGTGAGACTGAGATAGTTCTTGCGGTTGGTGGCTCT 1140
Qy 1141 GCTTAAGAGAGGTTGCTTAAGCTGTTTACGAGATGTTGTTTATGTTAGTTAGGCGCTGC 1200
Db 1141 GCTTAAGAGAGGTTGCTTAAGCTGTTTACGAGATGTTGTTTATGTTAGTTAGGCGCTGC 1200
Qy 1201 GTTTCCTGATGTCTCGAGGAGATGCTATGCAATGGATCCACTGATTTTATTAATCTGTTA 1260
Db 1201 GTTTCCTGATGTCTCGAGGAGATGCTATGCAATGGATCCACTGATTTTATTAATCTGTTA 1260
Qy 1261 TGAATTTGTGAGAGAGCTTTGAAGCTTTTACAGTATGTTGACTTGCTTTGGTAAATTTG 1320
Db 1261 TGAATTTGTGAGAGAGCTTTGAAGCTTTTACAGTATGTTGACTTGCTTTGGTAAATTTG 1320
Qy 1321 ACGAGCGTTGGCTTTATTAAGAACTTTCTGATTTGATCTTTGTTATGAGCTTGTTA 1380
Db 1321 ACGAGCGTTGGCTTTATTAAGAACTTTCTGATTTGATCTTTGTTATGAGCTTGTTA 1380
Qy 1381 GGAAGAGAGAGCAAGTACCTTTCGACCGGATTTACGTGACAAATTAAGTGAAGCTTTGGA 1440
Db 1381 GGAAGAGAGAGCAAGTACCTTTCGACCGGATTTACGTGACAAATTAAGTGAAGCTTTGGA 1440
Qy 1441 AGAGATCACTCGCGCTTATGCTTTCGAGCTTACCTTGCTTACCGCTTGCTGATGATTAAGC 1500
Db 1441 AGAGATCACTCGCGCTTATGCTTTCGAGCTTACCTTGCTTACCGCTTGCTGATGATTAAGC 1500
Qy 1501 TGGGAAAAGACTAAATGCTTTAAGCGGTGCGGAAATTTTGTGCTGTTGAGAGAGG 1560
Db 1501 TGGGAAAAGACTAAATGCTTTAAGCGGTGCGGAAATTTTGTGCTGTTGAGAGAGG 1560
Qy 1561 TGAAGCATCAGCTCTTGTGGGGGTTTGAACCGGTGAAGAGTTTATGATGAGGGCTTTT 1620
Db 1561 TGAAGCATCAGCTCTTGTGGGGGTTTGAACCGGTGAAGAGTTTATGATGAGGGCTTTT 1620
Qy 1621 ACGAATGACAGCTGCTGAGACAGATTAACAGTTTATGATCTTTTATTAATTTCTTTAGCA 1680
Db 1621 ACGAATGACAGCTGCTGAGACAGATTAACAGTTTATGATCTTTTATTAATTTCTTTAGCA 1680

|||||
Db 1621 ATGAATGACAGCTGCTGAGACAGATTAACAGTTTATGATCTTTTATTAATTTCTTTAGCA 1680
Qy 1681 TGAATTAATCTTATGAGTTTCTCATTTTAAATGATGTTGTGATGAGTTGATCTTTTGTGA 1740
Db 1681 TGAATTAATCTTATGAGTTTCTCATTTTAAATGATGTTGTGATGAGTTGATCTTTTGTGA 1740
Qy 1741 GCTAACCCCAAGCAATATTCAGACAGATCTTGAAGTTTACGAAGTTGACCTTGCTCTT 1800
Db 1741 GCTAACCCCAAGCAATATTCAGACAGATCTTGAAGTTTACGAAGTTGACCTTGCTCTT 1800
Qy 1801 GTGGCTCAAGCTTTTATGTTGATGAAGGACCACTTTTACAGGATGCTGATTAAGCAATTC 1860
Db 1801 GTGGCTCAAGCTTTTATGTTGATGAAGGACCACTTTTACAGGATGCTGATTAAGCAATTC 1860
Qy 1861 CAGCAACTTCAGACAGCTAAGTATGCTATGAGAGATTCCTGATGTTGATGATACA 1920
Db 1861 CAGCAACTTCAGACAGCTAAGTATGCTATGAGAGATTCCTGATGTTGATGATACA 1920
Qy 1921 CCGAATATATGGAGATGACCTTCGCTTGAAGAGGAGACTCTGTGCACTGCTTATAGGC 1980
Db 1921 CCGAATATATGGAGATGACCTTCGCTTGAAGAGGAGACTCTGTGCACTGCTTATAGGC 1980
Qy 1981 AAGTTGATGAATGCGGTATGTTGGAGAAATTCGATGAGAGATTCACAAATATAGGAAT 2040
Db 1981 AAGTTGATGAATGCGGTATGTTGGAGAAATTCGATGAGAGATTCACAAATATAGGAAT 2040
Qy 2041 CCAGCTATGTTGAGAGTTGTTTGGAGAAATTCGATGAGAGATTCGATGAGAGATTCGAT 2100
Db 2041 CCAGCTATGTTGAGAGTTGTTTGGAGAAATTCGATGAGAGATTCGATGAGAGATTCGAT 2100
Qy 2101 GGAATATGAAATTTGTTGGAACCTGCTGTTGGAAGGAGGTTCTTCTAGTTTCAAGAC 2160
Db 2101 GGAATATGAAATTTGTTGGAACCTGCTGTTGGAAGGAGGTTCTTCTAGTTTCAAGAC 2160
Qy 2161 ACCAATATGAAATTTTAACTCGGGGACTACTATGATGATCTTATGTTTGAATAC 2220
Db 2161 ACCAATATGAAATTTTAACTCGGGGACTACTATGATGATCTTATGTTTGAATAC 2220
Qy 2221 TTGGAAGAAGTGAAGTGAAGTTCAGAGGTTCTCTTATGAGCTGCTGCAATATGAGCAAGG 2280
Db 2221 TTGGAAGAAGTGAAGTGAAGTTCAGAGGTTCTCTTATGAGCTGCTGCAATATGAGCAAGG 2280
Qy 2281 ATTGAGCCGAGACATGTAAGGCTGATGATGAGCTGAGCTGAGCAAGAGTTTCTCTCC 2340
Db 2281 ATTGAGCCGAGACATGTAAGGCTGATGATGAGCTGAGCTGAGCAAGAGTTTCTCTCC 2340
Qy 2341 CGCTATACAGATTAAGAACTCGGCTGAACCCAGAGATGTCAGAGACAGGTTTATGATGA 2400
Db 2341 CGCTATACAGATTAAGAACTCGGCTGAACCCAGAGATGTCAGAGACAGGTTTATGATGA 2400
Qy 2401 GATCCTGTTGTAACATGTAAGGCTGATGATGAGCTGAGCTGTTTATTTGCAAGAGCT 2460
Db 2401 GATCCTGTTGTAACATGTAAGGCTGATGATGAGCTGAGCTGTTTATTTGCAAGAGCT 2460
Qy 2461 GTAAAGCTCTGTAAGAACTTTGAAACTAATGATTAAGCAATTCGAGCTGGGCTCAGAG 2520
Db 2461 GTAAAGCTCTGTAAGAACTTTGAAACTAATGATTAAGCAATTCGAGCTGGGCTCAGAG 2520
Qy 2521 AGTAGCGTTGATGAATCTAATGTTGAATGTCGTTGAGATGATGTTTAAAGAGAGCAAGT 2580
Db 2521 AGTAGCGTTGATGAATCTAATGTTGAATGTCGTTGAGATGATGTTTAAAGAGAGCAAGT 2580
Qy 2581 GTGAAGATCTTACGCTGCTGCTGAGCAATTTGAGCTGATTTCACTGTTCAAGCAGAAATAT 2640
Db 2581 GTGAAGATCTTACGCTGCTGCTGAGCAATTTGAGCTGATTTCACTGTTCAAGCAGAAATAT 2640
Qy 2641 TTTCTTAAAGAGCTCATCTTTTCAACGCAAGATATGTTTCTTATGAAATCTGAT 2700
Db 2641 TTTCTTAAAGAGCTCATCTTTTCAACGCAAGATATGTTTCTTATGAAATCTGAT 2700
Qy 2701 GTGCTATCAATAGATATGATTAATGATGCAATTTTCAATATCTGATGCTCAAAATA 2760
Db 2701 GTGCTATCAATAGATATGATTAATGATGCAATTTTCAATATCTGATGCTCAAAATA 2760
```

QY 2761 TCGTTGTTTGTGAGCTAAGAACATAGTTCCTTAATACATGTCCTCAAAAGTTGTAAC 2820
Db 2761 TCGTTGTTTGTGAGCTAAGAACATAGTTCCTTAATACATGTCCTCAAAAGTTGTAAC 2820
QY 2821 AAGATTACAAAGTTGCTGAGTAATTTCACTAATTAATGCTGTTGAATTTTGTATCAA 2880
Db 2821 AAGATTACAAAGTTGCTGAGTAATTTCACTAATTAATGCTGTTGAATTTTGTATCAA 2880
QY 2881 CTGTGACAGAAATGTAATTTCACTCAACATTTCTGTTAAGTAATGTAAGATTAG 2940
Db 2881 CTGTGACAGAAATGTAATTTCACTCAACATTTCTGTTAAGTAATGTAAGATTAG 2940
QY 2941 AGATTGCTTAAAGTGTGCTTGTGCTCAACTTTTCTTCTGATTTTCTTTGATTT 3000
Db 2941 AGATTGCTTAAAGTGTGCTTGTGCTCAACTTTTCTTCTGATTTTCTTTGATTT 3000
QY 3001 AGGTCATCTAGAGCTGACATTCAGAGCACTTCCAGAAATGATGCTAGAGACTGACA 3060
Db 3001 AGGTCATCTAGAGCTGACATTCAGAGCACTTCCAGAAATGATGCTAGAGACTGACA 3060
QY 3061 GAATTAATGATCCAAAGTGGAGAAATTAAGTCTGAGCTTTTGGGCTGATCAACCGAT 3120
Db 3061 GAATTAATGATCCAAAGTGGAGAAATTAAGTCTGAGCTTTTGGGCTGATCAACCGAT 3120
QY 3121 AGAAATGTTACAGAGCTGAGGGAATTAATCTACAAATTCATCAATGTTGTAAGAACTGT 3180
Db 3121 AGAAATGTTACAGAGCTGAGGGAATTAATCTACAAATTCATCAATGTTGTAAGAACTGT 3180
QY 3181 TGACATGATTAATGATCTGAGTCTGTTGTTGATTTCTGTTAATTAATGATTTGATGGCC 3240
Db 3181 TGACATGATTAATGATCTGAGTCTGTTGTTGATTTCTGTTAATTAATGATTTGATGGCC 3240
QY 3241 GAATGCTGAAGATTGGACTGACAGAGCACTGAACTGCGCAGCTTGGGTTGATTTAG 3300
Db 3241 GAATGCTGAAGATTGGACTGACAGAGCACTGAACTGCGCAGCTTGGGTTGATTTAG 3300
QY 3301 ATTATACACTGTTGAACTATCTGTTGACAGTGTGACAGCTCTACAGATGGAACCCGTG 3360
Db 3301 ATTATACACTGTTGAACTATCTGTTGACAGTGTGACAGCTCTACAGATGGAACCCGTG 3360
QY 3361 CTCTGCTGGAAGCACTCTGAGAGAGTCTGTTCTGTAATGATTTGTTGATCCAGAA 3420
Db 3361 CTCTGCTGGAAGCACTCTGAGAGAGTCTGTTCTGTAATGATTTGTTGATCCAGAA 3420
QY 3421 ACAATGCTAGTGTGGAAGCTGACACCTACACAAAGATAGAAATTTCTGCTGCAAGTAC 3480
Db 3421 ACAATGCTAGTGTGGAAGCTGACACCTACACAAAGATAGAAATTTCTGCTGCAAGTAC 3480
QY 3481 GGTGGAATGCTAGTGTGGAAGCTCTGTTCTGTCATCAATTAATGATGATGATGCT 3540
Db 3481 GGTGGAATGCTAGTGTGGAAGCTCTGTTCTGTCATCAATTAATGATGATGATGCT 3540
QY 3541 GAGCTTGGCAAGTCTCTTGTGTTGTAATTTCTCTCTAAGTAAGTGTATTAAGA 3600
Db 3541 GAGCTTGGCAAGTCTCTTGTGTTGTAATTTCTCTCTAAGTAAGTGTATTAAGA 3600
QY 3601 ACACAAAATTAATGATTTGTCGACACCCCTTTCTTGAATCAATTAATTAAG 3660
Db 3601 ACACAAAATTAATGATTTGTCGACACCCCTTTCTTGAATCAATTAATTAAG 3660
QY 3661 GCTACAA 3667
Db 3661 GCTACAA 3667

RESULT 3

US-10-739-930-227

; Sequence 227, Application US/10739930

; Publication No. US20040216190A1

; GENERAL INFORMATION:

; APPLICANT: Kovalic, David K.

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH

; TITLE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10/739,930
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 11088
; SEQ ID NO 227
; LENGTH: 2679
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURES:
; OTHER INFORMATION: Clone ID: ARATH-23APR03-CLUSTER13643_1
US-10-739-930-227

Query Match 53.7%; Score 1968.4; DB 8; Length 2679;
Best Local Similarity 80.5%; Pred. No. 0;
Matches 2657; Conservative 0; Mismatches 11; Indels 632; Gaps 5;

QY 368 GATTTAATTTTACTACTCAAAATCAAAATTCCTTAAGCCCTAGACGACCAACAGTCTC 427
Db 2 GATTTAATTTTACTACTCAAAATCAAAATTCCTTAAGCCCTAGACGACCAACAGTCTC 61
QY 428 TTCAATATGTAAGACAGAAAGTTTGTAGTACCTTAAGACACTCCATGTAAG 487
Db 62 TTCAATATGTAAGACAGAAAGTTTGTAGTACCTTAAGACACTCCATGTAAG 121
QY 488 CTGTGAGTCACTGCGGCAATTTGCTCTCCCAATTAATGCGATTAACACCGGCA 547
Db 122 CTGTGAGTCACTGCGGCAATTTGCTCTCCCAATTAATGCGATTAACACCGGCA 181
QY 548 CGACAAAGCTCCGAGTACCAACACACTCTCAACATCTGCTCCGCGCAAAATGG 607
Db 182 CGACAAAGCTCCGAGTACCAACACACTCTCAACATCTGCTCCGCGCAAAATGG 241
QY 608 CCGACCGCTCTCTCCGCAATTTCAATTCACCTCGATTCCTCTCTCTCTGCGCA 667
Db 242 CCGACCGCTCTCTCTCCGCAATTTCAATTCACCTCGATTCCTCTCTCTCTGCGCA 301
QY 668 CCGCACCAACCAACCGCACTCTGCTCTCTGCAACATCTAATGATGTCGCGAACGC 727
Db 302 CCGCACCAACCAACCGCACTCTGCTCTCTGCAACATCTAATGATGTCGCGAACGC 361
QY 728 ACGTCCCATCCCATTTGATTTTCAACAGTATTAAGAGCTCAACATTTCTTAACCG 787
Db 362 ACGTCCCATCCCATTTGATTTTCAACAGTATTAAGAGCTCAACATTTCTTAACCG 421
QY 788 ATGAATCAGAAAGCATGAGCTAGGTTTGAACCGCGCAATTCGTTCAAGC 847
Db 422 ATGAATCAGAAAGCATGAGCTAGGTTTGAACCGCGCAATTCGTTCAAGC 481
QY 848 ACGACGCTTAAATGACCGGAGACAGATCTTCAAGCTGTTGCAAACTCTGCTAATC 907
Db 482 ACGACGCTTAAATGACCGGAGACAGATCTTCAAGCTGTTGCAAACTCTGCTAATC 541
QY 908 CTCGCTCAGAAAGATGATCATATGAGTCTTCTTGAATGAAGAGACTACAGTCA 967
Db 542 CTCGCTCAGAAAGATGATCATATGAGTCTTCTTGAATGAAGAGACTACAGTCA 601
QY 968 CTGATGTTCTTGGGATTAAGTAAATTTGATTTGGAATTAATTAAGTTCTTCTGTTAA 1027
Db 602 CTGATGTTCTTGGAT----- 618
QY 1028 TTTGATGATTTGATTAAGAAAGAACTTTATCTAGTAAGGTTCTGAGGCTCTCTGT 1087
Db 619 -----AAGGTTCTGAGTCTCTGT 639
QY 1088 GTATTGCAAGAGTGTGAGACTGAGATAGTCTTCTGAGGTTGTGAGGCTCTGTTAAG 1147
Db 640 GTATTGCAAGAGTGTGAGACTGAGATAGTCTTCTGAGGTTGTGAGGCTCTGTTAAG 699
QY 1148 GAGAGGTTGCTTAAGTGTGTTAAGCAAGATGAGTGTGTTAAGTGTGAGGCTCTGCTCTC 1207
Db 700 GAGAGGTTGCTTAAGTGTGTTAAGCAAGATGAGTGTGTTAAGTGTGAGGCTCTGCTCTC 759

QY	1208	GATGCTTCGAGGAGNAGCTAATGGCAATTGGAATCCACTGATTTTATATACGCTAATGAGTTT	1267
Db	760	GATGCTTCGAGGAGNAGCTAATGGCAATTGGAATCCACTGATTTTATATACGCTAATGAGTTT	819
QY	1268	GTTGAGGAAGCTTTGAAGCTTTTACAGAGTAGTTGACTTGCTTGTGTAATTTGACGACGC	1327
Db	820	GTTGAGGAAGCTTTGAAGCTTTTACAGAGTAGTTGACTTGCTTGTGTAATTTGACGACGC	844
QY	1328	TTGGCTTTATTAAGAACTTTCTTGATTTGATACCTTGTGTTATTTGAGCTCTTGCTGACGAGAA	1387
Db	845	-----AGAGAGAA	852
QY	1388	GGAGCAATGACCTTGGCAACGGGATTTTACATGCAAAATTTGATGACCTTGGAGAGATC	1447
Db	853	GGAGCAATGACCTTGGCAACGGGATTTTACATGCAAAATTTGATGACCTTGGAGAGATC	912
QY	1448	ACTCCGCTTAATGCTTGGAGACTACTTGGCTTACCGCTTGATGATTAACGCTGCAAA	1507
Db	913	ACTCCGCTTAATGCTTGGAGACTACTTGGCTTACCGCTTGATGATTAACGCTGCAAA	972
QY	1508	AGACTAAATGCTTTAAGCGGTGTGCGGAATATTTTGTCGTGTTGAGAGAGGTGAGCA	1567
Db	973	AGACTAAATGCTTTAAGCGGTGTGCGGAATATTTTGTCGTGTTGAGAGAGGTGAGCA	1032
QY	1568	TCAGCTCTTGTGGGGGTTTGAACCCGTGAGAAAGTTATAGATGAGGCTTTTACGATG	1627
Db	1033	TCAGCTCTTGTGGGGGTTTGAACCCGTGAGAAAGTTATAGATGAGGCTTTTACGATG	1092
QY	1628	ACAGCTGCTGACGAGGTATACAGTTAGATACCTTTTAAATTTCTTACATGATATA	1687
Db	1093	ACAGCTGCTGAGC-----	1105
QY	1688	ACTTTAGGTTTCTCAATTTTAATGTAATGTGTGTGTGAGGTGATCTTTTGTGACCTAAC	1747
Db	1106	-----AGTTGATCTTTTGTGACCTAAC	1129
QY	1748	CAAGCAATTTTCACAGACATCTTGAAGTTTATAGAAAGTTGCACTTGCTCTTGTCGTC	1807
Db	1130	CAAGCAATTTTCACAGACATCTTGAAGTTTATAGAAAGTTGCACTTGCTCTTGTCGTC	1189
QY	1808	AAGCTTTATTTGTTAAGAACCAACCTTTTACAGAGTGTGATAGCAATTTCCAGCAAC	1867
Db	1130	AAGCTTTATTTGTTAAGAACCAACCTTTTACAGAGTGTGATAGCAATTTCCAGCAAC	1249
QY	1868	TTCAGCAGGCTAAGGTAAATGGCTATGAGATTCCTGCGAGTGTGTATGATACACGGAATA	1927
Db	1250	TTCAGCAGGCTAAGGTAAATGGCTATGAGATTCCTGCGAGTGTGTATGATACACGGAATA	1309
QY	1928	ATTGGGAATATGACTTCGCTCTAGAAAGGGAATCTGTGCACTGCTTATAGCCAAAGTTG	1987
Db	1310	ATTGGGAATATGACTTCGCTCTAGAAAGGGAATCTGTGCACTGCTTATAGCCAAAGTTG	1369
QY	1988	ATGAATGCGGATGTGTGTGGGCTTAGAAGTGAGAGATTCAAAATATGGAATCAAGCTA	2047
Db	1370	ATGAATGCGGATGTGTGTGGGCTTAGAAGTGAGAGATTCAAAATATGGAATCAAGCTA	1429
QY	2048	TTGTGAGATTGTTTTTGGAGAAATTCMAATCGTATGACATATGATATCTCCCTGACATAT	2107
Db	1430	TTGTGAGATTGTTTTTGGAGAAATTCMAATCGTATGACATATGATATCTCCCTGACATAT	1489
QY	2108	GCAAAATGTTGGAACCTGCTTGGCAGGGGTTTCTTCTTAGAGTTACAGACACCAAG	2167
Db	1490	GCAAAATGTTGGAACCTGCTTGGCAGGGGTTTCTTCTTAGAGTTACAGACACCAAG	1549
QY	2168	ATTAATAAATTTTAACTCGGGGACTATATATGATGATCTATGGTTTTGAGTTACTTGAAA	2227
Db	1550	ATTAATAAATTTTAACTCGGGGACTATATATGATGATCTATGGTTTTGAGTTACTTGAAA	1609
QY	2228	GAGTGGAGTAGTTTACAGGTTTCTCTTTAGCTGCTGCAACTATGCAAGATTTGGAG	2287
Db	1610	GAGTGGAGTAGTTTACAGGTTTCTCTTTAGCTGCTGCAACTATGCAAGATTTGGAG	1669
QY	2288	CCGACATATGTGAAGCTATGCTATGACAGGCACTGCAAGAAAGTTTTTCTTCCCGCTATA	2347

Db	1670	CCGAGCACTGTAAAGCTAGTGTATGACAGGCACTGCAGAAAGTTTTTCTTCCCGCTATA	1729
Qy	2348	CAGATAGAAACTCGGCTGAACCCAGAGATGTGCAGAGACAGCTGTTAGTGTAGATCTTG	2407
Db	1730	CAGATAGAAACTCGGCTGAACCCAGAGATGTGCAGAGACAGCTGTTAGTGTAGATCTTG	1789
Qy	2408	TTGTGTAAACAATGTAGCCGTGTATGTGTAGCCGTGTGTCTTTATTTGAGAAAGCTGTAAAGC	2467
Db	1790	TTGTGTAAACAATGTAGCCGTGTATGTGTAGCCGTGTGTCTTTATTTGAGAAAGCTGTAAAGC	1849
Qy	2468	CCTCTGAAAACTTTGAAAACTAATGATTAATGTCAATTCGACTGGGAGTCTCAGAGATAGCG	2527
Db	1850	CCTCTGAAAACTTTGAAAACTAATGATTAATGTCAATTCGACTGGGAGTCTCAGAGATAGCG	1909
Qy	2528	TTGATGAAACTA CTGTGAAATGTCGCTTGCTGATATGTTAAAGAGGCAAGTGTGAAGA	2587
Db	1910	TTGATGAAACTA CTGTGAAATGTCGCTTGCTGATATGTTAAAGAGGCAAGTGTGAAGA	1969
Qy	2588	TCCTTAGCTGCTGTGTGTGGCAATTGGACTGTATTCACCTGTCCAGCCGGAATATTTCTTA	2647
Db	1970	TCCTTAGCTGCTGTGTGTGGCAATTGGACTGTATTCACCTGTCCAGCCGGAATATTTCTTA	2029
Qy	2648	AAAGCAGCTCATCTTTCAACGCAAGATATGATTTCTCTATGGAATCGATGTGCTTA	2707
Db	2030	AAAGCAGCTCATCTTTCAACGCAAGATATGATTTCTCTATGGAATCGATGTGCTTA	2089
Qy	2708	CCATAGTATGATTAATAATGATGCAATTTTCATATATCTGCATTTGCTCAAAATATGCTTGT	2767
Db	2090	CCA-----	2092
Qy	2768	TTTGTGAGCTAAGAACATATGTTCCACTTAATACATGTCCCAAAAGTTGTACCAAGATTA	2827
Db	2093	-----	2092
Qy	2828	ACAAATTGCTAGTAATAATTTCACTAATATGCTGTGAATTTTTTGATCAAACTGTAGA	2887
Db	2093	-----	2092
Qy	2888	CAGAAATGTAAATTTCACTCTCAACATTTCTGTTTAGAAATACGTAGATTAGAGATTGC	2947
Db	2093	-----	2092
Qy	2948	CTTAATGTGGCTTTGTCCAACTTTTCTTTCTTGATTTTTTCTTTTCGATTTAAGGTC	3007
Db	2093	-----TAGGGTCA	2100
Qy	3008	GTCAGAGCTGACGATTCAGAGACCTTCCAGAAATGATGCTAGAGCTGCAGAGAAATATA	3067
Db	2101	GTCAGAGCTGACGATTCAGAGACCTTCCAGAAATGATGCTAGAGCTGCAGAGAAATATA	2160
Qy	3068	GTCATCAAGTGCAGAGAAATTAAGTCTCTGGCTTTTGAGCTGTATCCAGCATAGAAATG	3127
Db	2161	GTCATCAAGTGCAGAGAAATTAAGTCTCTGGCTTTTGAGCTGTATCCAGCATAGAAATG	2220
Qy	3128	TTACCAAGCTGAGGGAATTAATCTTACAAATTCATCAATTTGTGTAAAACTGTGGACAT	3187
Db	2221	TTACCAAG-----	2227
Qy	3188	GATTATAGTCTGTGCTTGTGTTGATTCGTATATTATAGTTTTTGATGGGGAATGCT	3247
Db	2228	-----AGGTTTTTGATGGGGAATGCT	2249
Qy	3248	GAAGATTTGGA CTGACAGAGCAGCTGAAACTCGCAGCTTGAGGTGTTATGATTTATAC	3307
Db	2250	GAAGATTTGGA CTGACAGAGCAGCTGAAACTCGCAGCTTGAGGTGTTATGATTTATAC	2309
Qy	3308	ACTGTTGAAACTATCTGTTGACAGTGTGA CAGTCTCAGAGATGAAACCCGTGCTCTGCT	3367
Db	2310	ACTGTTGAAACTATCTGTTGACAGTGTGA CAGTCTCAGAGATGAAACCCGTGCTCTGCT	2369
Qy	3368	GGAACCACTCTGAGAGAGTCTGCTTGTATCTGATTTTGTTTCAATCCAGAAACATATGC	3427

Db	2370	GGAAGCAACTGGAGGAGCTGCTGGTCATCTGATTTGGTTATCCAGAAAAACAATGC	2429
Qy	3428	TACTGANGTCGAACCTACACAAACAGATACGAAGTTTTCTGGTCCAAAGTCAGGGTGGAA	3488
Db	2430	TACTGAATGTCAAACTACACAAACAGATACGAAGTTTTCTGGTCCAAAGTCAGGGTGGAA	2489
Qy	3488	AATCACTGAAGGCTCTGTTCTTGATCATATATATACATATGATAGTACAGTCTGAGCTTG	3547
Db	2490	AATCACTGAAGGCTCTGTTCTTGATCATATATATACATATGATAGTACAGTCTGAGCTTG	2549
Qy	3548	CGAGATTCCTCTTGTCTGTAAATTCCTCTCTCTAAAGTTAGTGTATTAATGAACACAA	3607
Db	2550	CGAGATTCCTCTTGTGTAAATTCCTCTCTCTCTAAAGTTAGTGTATTAATGAACACAA	2609
Qy	3608	AAAATTACGTTCTTGGCACACCCTTTTCCTGATCTAAACTATATACATAAGGGCTCAAA	3667
Db	2610	AAAATTACGTTCTTGGCACAAAAAAATTCCTTGATCTAAACTATATACATAAGGGCTCAAA	2669

```

RESULT 4
US-10-600-070-130
: Sequence 130, Application US/10600070
: Publication No. US20040139500A1
: GENERAL INFORMATION:
: APPLICANT: Osteeryoung, Katherine W.
: APPLICANT: Viltha, Stanislaw
: APPLICANT: Koksharova, Olga A.
: APPLICANT: Gao, Hongbo
: TITLE OF INVENTION: Placid Division and Related Genes and Proteins, and Methods of
: TITLE OF INVENTION: Use
: FILE REFERENCE: MSU-08153
: CURRENT APPLICATION NUMBER: US/10/600,070
: CURRENT FILING DATE: 2003-06-20
: NUMBER OF SEQ ID NOS: 206
: SOFTWARE: PatentIn version 3.2
: SEQ ID NO 130
: LENGTH: 2637
: TYPE: DNA
: ORGANISM: Arabidopsis thaliana
: US-10-600-070-130

```

Query Match	Best Local Similarity	52.7%; Score 1931;	DB 7,	Length 2637;	
		80.4%; Pred. No. 0;			
Matches 2616;	Conservative	0;	Mismatches	5; Indels 633; Gaps 5	
QY	368	GAATTAACCTTA	TACTACTCAAAAT	CAAAATTCATTAACCTTGAACGACCAACAGCTCTC	427
Db	1	GATTTA	CTTATCTACTCAAAAT	CAAAATTCATTAACCTTGAACGACCAACAGCTCTC	60
QY	428	TTCAATATGTAAA	CAGAACAAAGTTTTTGATAGTACCTTA	AAAAAGACACTCCCATGGAAG	487
Db	61	TTCAATATGTAAA	CAGAACAAAGTTTTTGATAGTACCTTA	AAAAAGACACTCCCATGGAAG	120
QY	488	CTCTGAGTCACGT	CGGCAATTGGTCTCTCCCATTC	CAATTAATGCGAATTAACCGGCGCA	547
Db	121	CTCTGAGTCACGT	GGCAATGGTCTCTCCCATTC	CAATTAATGCGAATTAACCGGCGCA	180
QY	548	CGACAAAGCT	CCGACGTACCAACACTCTTA	CAACTATCTGCTCCGCAAGCAATAGGG	607
Db	181	CGACAAAGCT	CCGACGTACCAACACTCTTA	CAACTATCTGCTCCGCAAGCAATAGGG	240
QY	608	CCGACCGCTTCTC	TCGCGACTTCAATTTCACTCTCGAATTC	CTCCCTCTCTCTCTCTTCCGCA	667
Db	241	CCGACCGCTTCTC	TCGCGACTTCAATTTCACTCTCGAATTC	CTCCCTCTCTCTCTCTTCCGCA	300
QY	668	CCGCGACACCA	CCGCGCACTCTCGTCTCTCTGCA	CCACATTAATGATGTC	727
Db	301	CCGCGACACCA	CCGCGCACTCTCGTCTCTCTGCGCA	CCATTAATGATGTC	360
QY	728	ACGCGCCCATCC	CCATTAATTTTCAACGAGTATTAAGAGCTCA	ACAATTTCTTAACG	787
Db	361	ACGCGCCCATCC	CCATTAATTTTCAACGAGTATTAAGAGCTCA	ACAATTTCTTAACG	420

OY	788	ATGAAATCAGAAAGCAATTCGAACTGAGGGTTTGAAACCGCCGCAATTGGCTTCAAGC	847
Db	421	ATGAAATCAGAAAGCAATTCGAACTGAGGGTTTGAAACCGCCGCAATTGGCTTCAAGC	480
OY	848	ACGACGGCTTATCAGCCGAGAGACAGATTCTTCAAGCTGCTTGGAAACTGTCTAATC	907
Db	481	ACGACGGCTTATCAGCCGAGAGACAGATTCTTCAAGCTGCTTGGAAACTGTCTAATC	540
OY	908	CTCGGCTCAGAGAGAGTACATGAAAGTCTTCTTGATGATGAAGAACTACAGTCATCA	967
Db	541	CTCGGCTCAGAGAGAGTACATGAAAGTCTTCTTGATGATGAAGAACTACAGTCATCA	600
OY	968	CTGATGTTCCCTMGGAATAGGTAATTTCCGATTTTCGAAATATAAGTTTCTTCGTTTAA	1027
Db	601	CTGATGTTCCCTMGGAATAGGTAATTTCCGATTTTCGAAATATAAGTTTCTTCGTTTAA	617
OY	1028	TTTCATGAAATTCGATTAAGGAAGAACCTTTATCTAGTGAAGGTTCTCGGGGCTCTCTGT	1087
Db	618	TTTCATGAAATTCGATTAAGGAAGAACCTTTATCTAGTGAAGGTTCTCGGGGCTCTCTGT	638
OY	1088	GTAATTGCAAGAGGTGTGAGACTGAGATAGTCTTGCGGTTGTGTGAGGCTCTGCTTAAG	1147
Db	639	GTAATTGCAAGAGGTGTGAGACTGAGATAGTCTTGCGGTTGTGTGAGGCTCTGCTTAAG	698
OY	1148	GAGAGGTTGCTTAAGTGTGTTAAGCAAGATGTGTATTAGTATGCGCTTGGCTTCTC	1207
Db	699	GAGAGGTTGCTTAAGTGTGTTAAGCAAGATGTGTATTAGTATGCGCTTGGCTTCTC	758
OY	1208	GATGCTCGAAGGATGCTAATGGCAATTGGAATCCACTGATTTTATTACTGGTATAGATTT	1267
Db	759	GATGCTCGAAGGATGCTAATGGCAATTGGAATCCACTGATTTTATTACTGGTATAGATTT	818
OY	1268	GTTGAGGAAGCTTTGAAGCTTTTACAAGATGTTGAAGCTTCTTGTGTAATTTGACGACG	1327
Db	819	GTTGAGGAAGCTTTGAAGCTTTTACAAGATGTTGAAGCTTCTTGTGTAATTTGACGACG	843
OY	1328	TTTGCTTTAATAGAACTTTCTTGATTTTGATACCTTTGATTAAGTCTTGCTGTAGAGAA	1387
Db	844	TTTGCTTTAATAGAACTTTCTTGATTTTGATACCTTTGATTAAGTCTTGCTGTAGAGAA	851
OY	1388	GGAGCAGTAGACCTTGCACCGGAATTTACGTCACAAATTGATGAGACTTTGGAAGATCC	1447
Db	852	GGAGCAGTAGACCTTGCACCGGAATTTACGTCACAAATTGATGAGACTTTGGAAGATCC	911
OY	1448	ACTCCGGCTATGCTTGGAGCTACTTGTGCTTACCGCTTGATGATTTACGCTGCAGAA	1507
Db	912	ACTCCGGCTATGCTTGGAGCTACTTGTGCTTACCGCTTGATGATTTACGCTGCAGAA	971
OY	1508	AGACTAAATGCTTAAAGCGGTGTGCGGAATATTTTGTGCTGTGTGAGAGAGGTGAGCA	1567
Db	972	AGACTAAATGCTTAAAGCGGTGTGCGGAATATTTTGTGCTGTGTGAGAGAGGTGAGCA	1031
OY	1568	TCAGCTCTTGTTGGGGGTTTGAACCCGAGAAAGTTTATGAAATAGAGGGCTTTTACGAATG	1627
Db	1032	TCAGCTCTTGTTGGGGGTTTGAACCCGAGAAAGTTTATGAAATAGAGGGCTTTTACGAATG	1091
OY	1628	ACAGCTCTGAGAGGTAATCAGATTAGATACCTTTTAAATTTCTTTAGCANTATATA	1687
Db	1092	ACAGCTCTGAGG-----	1104
OY	1688	ACTTATAGTTTCTCATTTAATGTAATGTTGTGTGAGAGGTTGATCTTTTGTAGCTACCC	1747
Db	1105	ACTTATAGTTTCTCATTTAATGTAATGTTGTGTGAGAGGTTGATCTTTTGTAGCTACCC	1128
OY	1748	CAGCAATATTTCCAGCAGAGTCAATTTGAAAGTTTACGAAGTTGCACCTTGCTTGTGGCTC	1807
Db	1129	CAGCAATATTTCCAGCAGAGTCAATTTGAAAGTTTACGAAGTTGCACCTTGCTTGTGGCTC	1188
OY	1808	AAGCTTTATTTGGTAAAGACCACACTTTTACAGAGTGTGATTAAGCAATTCACAGAAC	1867
Db	1189	AAGCTTTATTTGGTAAAGACCACACTTTTACAGAGTGTGATTAAGCAATTCACAGAAC	1248
OY	1868	TTGAGCAGGCTAAGTAAAGTAAAGCTAATGAGAAATTTCTCGATGTTGTATGATACCGGAATA	1927

```
Db 1249 TTGACGAGCTAAGGTAATGCTATGAGAGATTCCTGCGATGTTGATATATACAGGAATA 1308
Qy 1298 ATGGGAGATATGACTTCGGTCTAGAAAGGGGACTCTGCGACTGCTTTATAGCAAAAGTTG 1987
Db 1309 ATGGGAGATATGACTTCGGTCTAGAAAGGGGACTCTGCGACTGCTTTATAGCAAAAGTTG 1368
Qy 1988 ATGAATGCCGATATGTTGGGCTTGAACAGTGAAGATTCACAAATATAGAAATCAGCTA 2047
Db 1369 ATGAATGCCGATATGTTGGGCTTGAACAGTGAAGATTCACAAATATAGAAATCAGCTA 1428
Qy 2048 TTGAGAGATTTGTTTGGAGAAATTCAAATCGTATGACAAATGATATCTCCCTGACATAT 2107
Db 1429 TTGAGAGATTTGTTTGGAGAAATTCAAATCGTATGACAAATGATATCTCCCTGACATAT 1488
Qy 2108 GCAAAATGTTGGAAAACCTGTTGGCAGGGGTTGTTCTTCTAGTTGCTGAGGACCAAAAG 2167
Db 1489 GCAAAATGTTGGAAAACCTGTTGGCAGGGGTTGTTCTTCTAGTTGCTGAGGACCAAAAG 1548
Qy 2168 ATAAAAAATTTAAACTCGGGGACTATATGATGATCTATGATGTTTGAAGTTACTTGAAA 2227
Db 1549 ATAAAAAATTTAAACTCGGGGACTATATGATGATCTATGATGTTTGAAGTTACTTGAAA 1608
Qy 2228 GAGTGAAGATGATTAAGAGTTCTCTTTAGCTGCTGCTGCACTATGCGAAGATTTGAG 2287
Db 1609 GAGTGAAGATGATTAAGAGTTCTCTTTAGCTGCTGCTGCACTATGCGAAGATTTGAG 1668
Qy 2288 CCGAGCATGTGAAAGCTAGTGTGCTATGACAGGCACTGACAGAAATTTTCTTCCGCTATA 2347
Db 1669 CCGAGCATGTGAAAGCTAGTGTGCTATGACAGGCACTGACAGAAATTTTCTTCCGCTATA 1728
Qy 2348 CAGATAGAACTCGGCTGAACTCCAGAGATGTCAGAGAGACAGTGTGATGATCTGTG 2407
Db 1729 CAGATAGAACTCGGCTGAACTCCAGAGATGTCAGAGAGACAGTGTGATGATCTGTG 1788
Qy 2408 TTGATGAACAATGTAGGCGGTGATGTTGAGCCCTGTTTATTTGCAAAAGCTGTAAAC 2467
Db 1789 TTGATGAACAATGTAGGCGGTGATGTTGAGCCCTGTTTATTTGCAAAAGCTGTAAAC 1848
Qy 2468 CCTCTGAAAACCTTTGAAAATAATGATATGCAATTCGAGCTGGGGTCTCAGAGAGTACG 2527
Db 1849 CCTCTGAAAACCTTTGAAAATAATGATATGCAATTCGAGCTGGGGTCTCAGAGAGTACG 1908
Qy 2528 TTGATGAACAATACTGTGTAAGATGTCGGTGTGATATGTTAAAGAGGCAAGTGTGAAG 2587
Db 1909 TTGATGAACAATACTGTGTAAGATGTCGGTGTGATATGTTAAAGAGGCAAGTGTGAAG 1968
Qy 2588 TCCATAGCTGCTGCTGCTGCAATTTGACATGATTTCACTGTGACGCCAGAAATTTTCTTA 2647
Db 1969 TCCATAGCTGCTGCTGCTGCAATTTGACATGATTTCACTGTGACGCCAGAAATTTTCTTA 2028
Qy 2648 AAAGCAGCTCATCTTTTCAACGCAAGATATGTTCTTTCTATGGAATCTGATGCGCTA 2707
Db 2029 AAAGCAGCTCATCTTTTCAACGCAAGATATGTTCTTTCTATGGAATCTGATGCGCTA 2088
Qy 2708 CCATAGTATATGATTAATGATGCAATTTTCAATATATCTGCAATTCCTCAAAATATGCTTGT 2767
Db 2089 CCA----- 2091
Qy 2768 TTTGAGACTAAGAACATATGTTCCCACTTAATATACATGTCACAAAGTTGTACCAAGATTA 2827
Db 2092 ----- 2091
Qy 2828 ACAAGTGTCTGATTAATTTTCACTAATATATGCTGCTGAATTTTGTATCAAACTGTAGA 2887
Db 2092 ----- 2091
Qy 2888 CAGAAATGTAATTTCACTCTCAACATTTCTGTTAGATTAAGATTAAGAGATTGCTC 2947
Db 2092 ----- 2091
Qy 2948 CTTAGTGTGGCTTTGCTCAACTTTTCTTCTTGAATTTTCTTTGATTTAGGCTCA 3007
Db 2092 ----- 2091
```

```
Db 2092 -----TAGGCTCA 2099
Qy 3008 GTGACAGCTGACGATTCAGAGACACTTCCAGATGATGATAGACTGACAGAAATATA 3067
Db 2100 GTGACAGCTGACGATTCAGAGACACTTCCAGATGATGATAGACTGACAGAAATATA 2159
Qy 3068 GTATCCAAAGTGGCAGAAAGATTAAGTCTCTGCTTTTGGGCTTGATCAACCCATATGAATG 3127
Db 2160 GTATCCAAAGTGGCAGAAAGATTAAGTCTCTGCTTTTGGGCTTGATCAACCCATATGAATG 2219
Qy 3128 TTACCAAGGTGAGGGAATTAATCTACATTCATCAATTTGTGTGAAGAACTGTTGACAT 3187
Db 2220 TTACCAAG----- 2226
Qy 3188 GATTATAGTCTGTGCTGCTGTTGATTCGTATTTATAGTTTGTGATGGCGAATGCT 3247
Db 2227 -----AGGTTTGTGATGGCGAATGCT 2248
Qy 3248 GAAGATTTGACCTGACAGAGACGCTGAACCTGCGACGCTTGGTTGTTATGATTTATAC 3307
Db 2249 GAAGATTTGACCTGACAGAGACGCTGAACCTGCGACGCTTGGTTGTTATGATTTATAC 2308
Qy 3308 ACTGTTGAACTATCTGTTGACAGTGTGACAGTCTGACAGATGGAACCGGTGCTGCT 3367
Db 2309 ACTGTTGAACTATCTGTTGACAGTGTGACAGTCTGACAGATGGAACCGGTGCTGCT 2368
Qy 3368 GGAAGCAACTCTGAGAGAGTCTGCTTGTCTATCTGATGATTTGTTGATTCAGAAAACAATGC 3427
Db 2369 GGAAGCAACTCTGAGAGAGTCTGCTTGTCTATGATGATTTGTTGATTCAGAAAACAATGC 2428
Qy 3428 TACTGATGTGAGAACCTTACACACAAATATGCAAGTTTCTGCTCCAAAGTCAAGGTGGA 3487
Db 2429 TACTGATGTGAGAACCTTACACACAAATATGCAAGTTTCTGCTCCAAAGTCAAGGTGGA 2488
Qy 3488 AATCAGTGAAGGCTCTGTTCTGTCATATATATATATATATGATGATGATGATGATGATG 3547
Db 2489 AATCAGTGAAGGCTCTGTTCTGTCATATATATATATATATATGATGATGATGATGATG 2548
Qy 3548 CGAATTCCTCTTGTCTGTAATTTCTCTCTAAGTTAGTGTATTAATGAACACAAA 3607
Db 2549 CGAATTCCTCTTGTGTGTAATTTCTCTCTAAGTTAGTGTATTAATGAACACAAA 2608
Qy 3608 AAAATTAACGTTTC 3620
Db 2609 AAAATTAACGTTTC 2621
```

```
RESULT 5
US-10-600-070-1
; Sequence 1, Application US/10600070
; Publication No. US20040139500A1
; GENERAL INFORMATION:
; APPLICANT: Oseireyoung, Katherine W.
; APPLICANT: Vitsha, Stanislaw
; APPLICANT: Koksheva, Olga A.
; APPLICANT: Gao, Hong
; TITLE OF INVENTION: Placid Division and Related Genes and Proteins, and Methods of
; FILE REFERENCE: MSU-08153
; CURRENT FILING DATE: 2003-06-20
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 2406
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-600-070-1
```

Query Match 47.0%; Score 1724; DB 7; Length 2406;
Best Local Similarity 79.2%; Pred. No. 0;
Matches 2406; Conservative 0; Mismatches 0; Indels 632; Gaps 5;

[illegible][illegible]

```
Db 1909 TTTCTTAAAGAGCTCATCTTTTCAACGACGAGTATGTTTCTTCTATGGAATCTGAT 1968
Qy 2701 GTGGCTACCATAGATGATGATTAATATGATGCAATTTTCATATATCTGATCTCATAATA 2760
Db 1969 GTGGCTACCA----- 1978
Qy 2761 TGCTTTGTTTGTGAGCTAAGAACATATGTTCCACTTAATACATGTCACAAAAGTTGACC 2820
Db 1979 ----- 1978
Qy 2821 AAGATTACAAGTTGCTGATGATTAATTTTCACTAATATGCTGCTGAATTTTGTATCAA 2880
Db 1979 ----- 1978
Qy 2881 CTGTAGACAGAAATGTAATTTTCACTCTCAATTTCTGTTAGATAACGTAGGATTAAG 2940
Db 1979 ----- 1978
Qy 2941 AGATTGCTTATGTTGCTGCTTGTCTCAACTTTTCTTCTGATTTTCTTTTCTGATTT 3000
Db 1979 ----- 1979
Qy 3001 AGGCTACGTCAGAGCTGACGATTCAGAGCACTTCCAGAAATGATGCTAGAGCTGACGA 3060
Db 1980 AGGCTACGTCAGAGCTGACGATTCAGAGCACTTCCAGAAATGATGCTAGAGCTGACGA 2039
Qy 3061 GAATATAGTATCCAAAGTGGACGAAAGATTAAAGTCTGCTGCTTTTGGGCTGATCAACGAT 3120
Db 2040 GAATATAGTATCCAAAGTGGACGAAAGATTAAAGTCTGCTGCTTTTGGGCTGATCAACGAT 2099
Qy 3121 AGAAATGTTACCAAGCTGAGGGAATTAATCTCAATTCATCAATGTTGTGAATACTGT 3180
Db 2100 AGAAATGTTACCAAG----- 2113
Qy 3181 TGGAATATTAATGCTGAGTCTGCTGCTTGTGATCTGTTATTTATAGGTTTGTGATGGGC 3240
Db 2114 ----- 2128
Qy 3241 GAATGCTGAGAGATTGGAATGACGACAGAGAGCTGAAATCTGCGAGCTTGGTGGTTATG 3300
Db 2129 GAATGCTGAGAGATTGGAATGACGACAGAGAGCTGAAATCTGCGAGCTTGGTGGTTATG 2188
Qy 3301 ATTATACCTGTTGAAACTATCTGTTGACAGTGTACAGTCTCAGACAAATGGAACCGGTG 3360
Db 2189 ATTATACCTGTTGAAACTATCTGTTGACAGTGTACAGTCTCAGACAAATGGAACCGGTG 2248
Qy 3361 CTCTGCTGAGAGCACTCTGAGAGAGTCTGCTGCTATCTGATTTGTTGATCCAGAA 3420
Db 2249 CTCTGCTGAGAGCACTCTGAGAGAGTCTGCTGCTATCTGATTTGTTGATCCAGAA 2308
Qy 3421 ACAATGCTACTGATGTCAGAACTCTACACAAAGATTAAGAAATTTCTGTCGAAGTCAG 3480
Db 2309 ACAATGCTACTGATGTCAGAACTCTACACAAAGATTAAGAAATTTCTGTCGAAGTCAG 2368
Qy 3481 GGTGGAATAATCACTGAGAGGCTCTGTTCTTGCAATCAAAA 3518
Db 2369 GGTGGAATAATCACTGAGAGGCTCTGTTCTTGCAATCAAAA 2406

RESULT 6
US-10-600-070-9
; Sequence 9, Application US/10600070
; Publication No. US20040139500A1
; GENERAL INFORMATION:
; APPLICANT: Osteoryoung, Katherine W.
; APPLICANT: Vittha, Stanislaw
; APPLICANT: Kokeharova, Olga A.
; APPLICANT: Gao, Hong
; TITLE OF INVENTION: Placitd Division and Related Genes and Proteins, and Methods of
; FILE REFERENCE: MSU-09153
; CURRENT APPLICATION NUMBER: US/10/600,070
; CURRENT FILING DATE: 2003-06-20
```

```
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 9
; LENGTH: 2406
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-600-070-9

Query Match 46.9%; Score 1720.8; DB 7; Length 2406;
Best Local Similarity 79.1%; Pred. No. 0;
Matches 2404; Conservative 0; Mismatches 2; Indels 632; Gaps 5;

Qy 481 ATGGAAGCTCTGAGTCACTGCGGATGCTGCTCTCCCATTCCAATATGCGCATTAACA 540
Db 1 ATGGAAGCTCTGAGTCACTGCGGATGCTGCTCTCCCATTCCAATATGCGCATTAACA 60
Qy 541 CCGGCGAGCAAAAGCTCCGAGGACACAAACCTCTCAACTATCTGCTCGGACG 600
Db 61 CCGGCGAGCAAAAGCTCCGAGGACACAAACCTCTCAACTATCTGCTCGGACG 120
Qy 601 AAATGGCCGACGCTCTTCTCTCCGACTTCATTTCACTCCGATTCCTCTCTCC 660
Db 121 AAATGGCCGACGCTCTTCTCTCCGACTTCATTTCACTCCGATTCCTCTCTCC 180
Qy 661 TTGCGCAACCGCACCAACCAACGCACTCTGCTCTCTGCAACCATTAATGATGTC 720
Db 181 TTGCGCAACCGCACCAACCAACGCACTCTGCTCTCTGCAACCATTAATGATGTC 240
Qy 721 GAAGCGACGCTCCCATCCCATTTGATTTGACAGGATTAAGAGCTCAAAACATTT 780
Db 241 GAAGCGACGCTCCCATCCCATTTGATTTGACAGGATTAAGAGCTCAAAACATTT 300
Qy 781 TTAAACGATGGAATCAGAAAGCAATTCGAAGCTAGAGGTTTGAACCCGCGCAATTCGGT 840
Db 301 TTAAACGATGGAATCAGAAAGCAATTCGAAGCTAGAGGTTTGAACCCGCGCAATTCGGT 360
Qy 841 TTACGCGACGACGCTTTAATCAACGCGAGACAGATTTTCAAGCTCTTGGCAACTCTG 900
Db 361 TTACGCGACGACGCTTTAATCAACGCGAGACAGATTTTCAAGCTCTTGGCAACTCTG 420
Qy 901 TCTAATCTCGGCTAGAAAGAGTACAAATGAAGGCTCTTCTGATGATGAAGAGCTACA 960
Db 421 TCTAATCTCGGCTAGAAAGAGTACAAATGAAGGCTCTTCTGATGATGAAGAGCTACA 480
Qy 961 GTCATCACTGATGTTCTTGGGATTAAGTAATTCGATTTGGAATAAAGTTTCTTC 1020
Db 481 GTCATCACTGATGTTCTTGGGAT----- 504
Qy 1021 GTTTTAATTCATGATTAAGTAAGTAAGAGAACTTTAATCTAGTGAAGTTCTTGGGCG 1080
Db 505 ----- 518
Qy 1081 TCTCTGATTAATGCAAGAGTGTGAGACTGATGTTCTTGGGTTGTGAGGCTCT 1140
Db 519 TCTCTGATTAATGCAAGAGTGTGAGACTGATGTTCTTGGGTTGTGAGGCTCT 578
Qy 1141 GCTTAAGAGAGTGTGCTTAAGTCTTTAAGCAAGATGAGTTTAACTTAAGCGCTTGC 1200
Db 579 GCTTAAGAGAGTGTGCTTAAGTCTTTAAGCAAGATGAGTTTAACTTAAGCGCTTGC 638
Qy 1201 GTTCTCTGATGCTCGAGGAGTGTATGATGATGATGATGATGATGATGATGATGATGAT 1260
Db 639 GTTCTCTGATGCTCGAGGAGTGTATGATGATGATGATGATGATGATGATGATGATGAT 698
Qy 1261 TGAATTTGAGAGAACTTTGAAGCTTTTAAACAGATGATGATGATGATGATGATGATGAT 1320
Db 699 TGAATTTGAGAGAACTTTGAAGCTTTTAAAC----- 730
Qy 1321 ACGAGCGTGGCTTTAAGAACTTTGATTTGATTAATCTTGTATGATGATGATGATGAT 1380
Db 731 ----- 731
Qy 1381 GGAGAGAGAGCAAGTATGCTTGCACCGGATTTAAGTGCACAAATTAATGAGACTTTGGA 1440
```

[illegible]

Db	1729	GTAAAGACCCTGTGAAAACCTTTGAAACCTAATGATTATTCGAAATTTGAGCTGGGGTCTCAG	1788
Oy	2521	AGTAGCGTGTATGAACTACTGTTGAAATGTCGGTGTCTATATGTTAAAGAGCGAAGT	2580
Db	1789	AGTAGCGTGTATGAACTACTGTTGAAATGTCGGTGTCTATATGTTAAAGAGCGAAGT	1848
Oy	2581	GTGAAGATTCCTAGCTGTGGTGTGGCAATTGAGCTGATTTCACTGTTCAAGCGAAGTAT	2640
Db	1849	GTGAAGATTCCTAGCTGTGGTGTGGCAATTGAGCTGATTTCACTGTTCAAGCGAAGTAT	1908
Oy	2641	TTTCTTTAAAGACAGCTCATCTTTTCAACGCAAGATATGAGTTCTTCTATGAAATCTGAT	2700
Db	1909	TTTCTTTAAAGACAGCTCATCTTTTCAACGCAAGATATGAGTTCTTCTATGAAATCTGAT	1968
Oy	2701	GTCGCTACCAATAGGTATGATTAATGAATGCAATTTTCATATATCTGCATGCTCCAAATA	2760
Db	1969	GTCGCTACCA-----	1978
Oy	2761	TGCTTTGTTTGTAGCTTAAGAACATAGTTCACCTTATATCATGTCCCAAAAGTTGACC	2820
Db	1979	-----	1978
Oy	2821	AAGATTAACAAGTTTGTGAGTAATTTCACTAATATATGCTGTGAATTTTGTGATCAAA	2880
Db	1979	-----	1978
Oy	2881	CTGTAGACAGAAATGTAAATTTCACTCTCAACATTTCTGTTTGAATACGTAGATTAG	2940
Db	1979	-----	1978
Oy	2941	AGATTGCTTATGTGTGGCTTTGTTCCAACTTTTCTTTCCTTGAATTTTTCCTTTGCATTT	3000
Db	1979	-----	1979
Oy	3001	AGGGTCAGTCAAGAGCTGACGATTCAAGAGCACTTCCAGAAATGGATGCTAGGACTGACA	3060
Db	1980	AGGGTCAGTCAAGAGCTGACGATTCAAGAGCACTTCCAGAAATGGATGCTAGGACTGACA	2039
Oy	3061	GAATATAGTATCCAAAGTGAGCAAGAAATTAAGTCTCTGGCTTTTGGGCTGATCACGCGAT	3120
Db	2040	GAATATAGTATCCAAAGTGAGCAAGAAATTAAGTCTCTGGCTTTTGGGCTGATCACGCGAT	2099
Oy	3121	AGAAATGTTTACAGAGGTGAGGAATTAATCTAACATTCATCAATTTGTGTAAAACTGT	3180
Db	2100	AGAAATGTTTACAG-----	2113
Oy	3181	TGAGCAGATTAATATAGTCTGTGGCTTGTTGATTCTGTTATTTATATAGTTTGGATGGCG	3240
Db	2114	-----	2128
Oy	3241	GAATGCTGAAGATTTTGGACTGACAGAGCAGCTGAAATCGCGCAGCTTGGGTTGGTTATG	3300
Db	2129	GAATGCTGAAGATTTTGGACTGACAGAGCAGCTGAAATCGCGCAGCTTGGGTTGGTTATG	2188
Oy	3301	ATTATACACTGTTGAAACTATCTGTGTGACAGTGTGACAGTCTGACAGATGAAACCCGTG	3360
Db	2189	ATTATACACTGTTGAAACTATCTGTGTGACAGTGTGACAGTCTGACAGATGAAACCCGTG	2248
Oy	3361	CTCTGTGTGAAAGCACTCTGAGAGAGTCTGCTTGTCTATCTGATTTGGTTCATCCAGAAA	3420
Db	2249	CTCTGTGTGAAAGCACTCTGAGAGAGTCTGCTTGTCTATCTGATTTGGTTCATCCAGAAA	2308
Oy	3421	ACATGTCTACTGTATGTACAGAACTTACACAAACAAGATACGAAGTTTCTGTGCCAAGTCAG	3480
Db	2309	ACATGTCTACTGTATGTACAGAACTTACACAAACAAGATACGAAGTTTCTGTGCCAAGTCAG	2368
Oy	3481	GGTGAAGAAATCACTGGAAGGCTCTGTTCTTTCGATCATATA 3518	
Db	2369	GGTGAAGAAATCACTGGAAGGCTCTGTTCTTTCGATCATATA 2406	

RESULT 7
US-10-600-070-128

Sequence 128, Application US/10600070
; Publication No. US20040139500A1
; GENERAL INFORMATION:
; APPLICANT: OsteoYoung, Katherine W.
; APPLICANT: Vicha, Stanislav
; APPLICANT: Koksharova, Olga A.
; APPLICANT: Gao, Hongo
; TITLE OF INVENTION: Plasmid Division and Related Genes and Proteins, and Methods of
; FILE REFERENCE: Use
; FILE REFERENCE: MSU-08153
; CURRENT APPLICATION NUMBER: US/10/600,070
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 128
; LENGTH: 2406
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-600-070-128

Query Match 46.8%; Score 1717.6; DB 7; Length 2406;
Best Local Similarity 79.1%; Pred. No. 0;
Matches 2402; Conservative 0; Mismatches 4; Indels 632; Gaps 5;

QY 481 ATGGAAGCTTGAGTCAAGTGGGCAATGGTCTCTCCCATTCGAATTATGCCGATTACCA 540
DB 1 ATGGAAGCTTGAGTCAAGTGGGCAATGGTCTCTCCCATTCGAATTATGCCGATTACCA 60
QY 541 CCGGGAAGCAAAAGCTCCGAGTGGAGCAAAACCTTACAACTATGCTCCGCGCAGC 600
DB 61 CCGGGAAGCAAAAGCTCCGAGTGGAGCAAAACCTTACAACTATGCTCCGCGCAGC 120
QY 601 AAATGGGCGGACCGCTCTCTCCGACTTCAATTCACCTCCGATTCCTCTCTCTCC 660
DB 121 AAATGGGCGGACCGCTCTCTCCGACTTCAATTCACCTCCGATTCCTCTCTCTCC 180
QY 661 TTCCGCAACCGCAACCAACCAACCGCACTCTCTCTCTCCGCAACATCTATTTGATGCTCC 720
DB 181 TTCCGCAACCGCAACCAACCAACCGCACTCTCTCTCTCCGCAACATCTATTTGATGCTCC 240
QY 721 GAAGCGCAAGTCCCATTCGATTTCTTACCAAGTATTAAGAGCTTCAAAACATTTTC 780
DB 241 GAAGCGCAAGTCCCATTCGATTTCTTACCAAGTATTAAGAGCTTCAAAACATTTTC 300
QY 781 TTAACCGATGGAATCAGAAGAGCATTCGAGCTAAGGTTTGAACACCGCCCAATTCGCT 840
DB 301 TTAACCGATGGAATCAGAAGAGCATTCGAGCTAAGGTTTGAACACCGCCCAATTCGCT 360
QY 841 TTCAGCGCAAGCTTTATATCAGCCCGGAGACAGATTTCTTCAAGCTGCTTGGAAACTCTG 900
DB 361 TTCAGCGCAAGCTTTATATCAGCCCGGAGACAGATTTCTTCAAGCTGCTTGGAAACTCTG 420
QY 901 TCTATCTCTCGGTCTAGAAAGAGATGATGAGTCTTCTTGAATGATGAAGAGCTTACA 960
DB 421 TCTATCTCTCGGTCTAGAAAGAGATGATGAGTCTTCTTGAATGATGAAGAGCTTACA 480
QY 961 GTCAATCATGATGTTCTTGGGATTAAGGTAATTTGATTTGGAATATATAAGTTTCTTC 1020
DB 481 GTCAATCATGATGTTCTTGGGAT----- 504
QY 1021 GTTTTAAATTTCAATGATGATTAAGAGAACTTTTATCTAGTGAAGGTTCTCGGGC 1080
DB 505 -----AAGTTTCTCGGTGC 518
QY 1081 TCTCTGTATTTGCAAGAGGTGATGAGTCTGATGCTCTCGGGTGGTGAAGGCTCT 1140
DB 519 TCTCTGTATTTGCAAGAGGTGATGAGTCTGATGCTCTCGGGTGGTGAAGGCTCT 578
QY 1141 GCTTAAGAGAGGTTGCTTAAGTCTTTAAGCAAGATGTTGTTTATGTTATGCGCTTGC 1200
DB 579 GCTTAAGAGAGGTTGCTTAAGTCTTTAAGCAAGATGTTGTTTATGTTATGCGCTTGC 638
QY 1201 GTTTCGTGATGTTCTGAGGAGTGTATGAGCATTTGATCCACCTGATTTTATTAATGTTA 1260

DB 639 GTTTCGTGATGTTCTGAGGAGTGTATGAGCATTTGATCCACCTGATTTTATTAATGTTA 698
QY 1261 TGAGTTTGTGAGGAGCTTTGAAGCTTTTACAGAGTATGATCTTGCTTGTGATTTTG 1320
DB 699 TGAGTTTGTGAGGAGCTTTGAGGCTTTTAC----- 730
QY 1321 ACAGAGCTTGCTTATTAAGAACTTTCTTGAATTTGATCTTTGATTTGATGCTTGTGA 1380
DB 731 -----A 731
QY 1381 GGAGGAAAGAGCAATGATGCTTGGACCGGATTTACGTGACCAATTTGATGAGATTTTGA 1440
DB 732 GGAGGAAAGAGCAATGATGCTTGGACCGGATTTACGTGACCAATTTGATGAGATTTTGA 791
QY 1441 AGAGATCACTCCGCGTATGTTCTTGGAGTACTTGGCTTACCGCTTGGTATGATTTACG 1500
DB 792 AGAGATCACTCCGCGTATGTTCTTGGAGTACTTGGCTTACCGCTTGGTATGATTTACG 851
QY 1501 TGCAGAAAGACTAATGTTTAAGCGGTGTGCGAATATTTTGTGTCTGTGAGAGAGG 1560
DB 852 TGCAGAAAGACTAATGTTTAAGCGGTGTGCGAATATTTTGTGTCTGTGAGAGAGG 911
QY 1561 TGGAGCATCAGCTTGTGTGGGGTTTGACCCGTGAGAACTTTATGAATGAGGCTTTT 1620
DB 912 TGGAGCATCAGCTTGTGTGGGGTTTGACCCGTGAGAACTTTATGAATGAGGCTTTT 971
QY 1621 ACAGATGACAGCTGCTGAGGAGTATACAGTTTATGATCTTTTAAATTTCTTACGA 1680
DB 972 ACAGATGACAGCTGCTGAG----- 991
QY 1681 TGATATTAATTTAGTTTCTCATTTTAAATGATGTTGTGTGAGTGTGATCTTTTGA 1740
DB 992 -----AGTTTGAATCTTTTGA 1008
QY 1741 GCTTACCCCAAGCAATATTTCCAGCAGAGTCAATTTGAATTTTACGAATTTGACATTTGCTCT 1800
DB 1009 GCTTACCCCAAGCAATATTTCCAGCAGAGTCAATTTGAATTTTACGAATTTGACATTTGCTCT 1068
QY 1801 GTGAGCTCAAGCTTTTATTTGTTAAGAAAGCCACACTTTTACAGAGTGTGATTAAGCAATTC 1860
DB 1069 GTGAGCTCAAGCTTTTATTTGTTAAGAAAGCCACACTTTTACAGAGTGTGATTAAGCAATTC 1128
QY 1861 CAGCACTTTCAGCAGGCTAAGTATGCTATGAGATTCCTGAGTGTGATGATGATACA 1920
DB 1129 CAGCACTTTCAGCAGGCTAAGTATGCTATGAGATTCCTGAGTGTGATGATGATACA 1188
QY 1921 CGGAATTAATTTGGAGATGACTTCCGTCTTGAAGAAAGGAGACTCTGTGACCTCTTATAGGC 1980
DB 1189 CGGAATTAATTTGGAGATGACTTCCGTCTTGAAGAAAGGAGACTCTGTGACCTCTTATAGGC 1248
QY 1981 AAAGTTGATGAATGCTGATGTTGGCTTGAACAGTGAAGATTTCAATATATAGGAAT 2040
DB 1249 AAAGTTGATGAATGCTGATGTTGGCTTGAACAGTGAAGATTTCAATATATAGGAAT 1308
QY 2041 CCAAGCTATTTGAGATTTGTTTGGAGATTTCAATCTGATGAACAATGATATCTTCCT 2100
DB 1309 CCAAGCTATTTGAGATTTGTTTGGAGATTTCAATCTGATGAACAATGATATCTTCCT 1368
QY 2101 GGAATATGAGAAATTTGTTGAGAACTGTTGGCAGGGGTTCCTTCTAGAGTTCAAGAC 2160
DB 1369 GGAATATGAGAAATTTGTTGAGAACTGTTGGCAGGGGTTCCTTCTAGAGTTCAAGAC 1428
QY 2161 ACCAATATTAATAATTTAATCTGGGGACTACTATGATGATCTTATGTTTGAATTTAC 2220
DB 1429 ACCAATATTAATAATTTAATCTGGGGACTACTATGATGATCTTATGTTTGAATTTAC 1488
QY 2221 TTGGAAGAGTGAAGTATGATTCAGGGTTCTCTTATGCTGCTGCTGCACTATGAGCAAG 2280
DB 1489 TTGGAAGAGTGAAGTATGATTCAGGGTTCTCTTATGCTGCTGCTGCACTATGAGCAAG 1548
QY 2281 ATTGAGCCGAGCATGTTGAAGGATGATGAGCACTGCAAGAAATTTTCTTCCTCC 2340

Db 1549 ATTGAGCCGAGATGTGAAGCTAGTCTATGACGACCTGACGAAGTTTCTTCC 1608
Qy 2341 CGCTATACAGATAGAACTCGCTGAACCCAGAGATGTGACAGACAGTGTAGTGA 2400
Db 1609 CGCTATACAGATAGAACTCGCTGAACCCAGAGATGTGACAGACAGTGTAGTGA 1668
Qy 2401 GATCCGTGTGTGATACAGATGAGCCGTGATGTGACCTGTGTCTTTATTGACAGAGCT 2460
Db 1669 GATCCGTGTGTGATACAGATGAGCCGTGATGTGACCTGTGTCTTTATTGACAGAGCT 1728
Qy 2461 GTAAAGACCTCTGAAAACCTTGAACCTAATGATTAATGACATTCGAGCTGGGCTGACAG 2520
Db 1729 GTAAAGACCTCTGAAAACCTTGAACCTAATGATTAATGACATTCGAGCTGGGCTGACAG 1788
Qy 2521 AGTAGCGTTGATGAAACTACTGTGGAATGTGCGTGTGATATGTGTAAGAGGACAGT 2580
Db 1789 AGTAGCGTTGATGAAACTACTGTGGAATGTGCGTGTGATATGTGTAAGAGGACAGT 1848
Qy 2581 GTGAAGATCTGAGCTGT 2640
Db 1849 GTGAAGATCTGAGCTGT 1908
Qy 2641 TTTCTTAAAGAGCTCATCTTTTCAAGCAGATATGTTCTTCTATGATCTGAT 2700
Db 1909 TTTCTTAAAGAGCTCATCTTTTCAAGCAGATATGTTCTTCTATGATCTGAT 1968
Qy 2701 GTGCTACCATGATGATATGATTAATGATGATTAATGATTAATGATTAATGATTAAT 2760
Db 1969 GTGCTACCATGATGATATGATTAATGATGATTAATGATTAATGATTAATGATTAAT 1978
Qy 2761 TGTCTGTTTGTGAGCTAAGAACATATGTTCCCATTAATATATGTTCCAAAGTTGTACC 2820
Db 1979 ----- 1978
Qy 2821 AAGATTAACAAGTTGCTGAGTAATTTCACTAATTAATGCTGCTGATTTTTCATCAA 2880
Db 1979 ----- 1978
Qy 2881 CTGTAGACAGAAATGTAATTTCACTGTCACATTCATGTTTGAATATGATGATTAAG 2940
Db 1979 ----- 1978
Qy 2941 AGATTGCTTATGATGTGCTGTTGTCCAACTTTTCTTCTGATTTTCTTCTGATTT 3000
Db 1979 ----- 1979
Qy 3001 AGGTCAGTCAGAGCTGACGATTCAGAGCACTTCCAGAAATGATGATGATGATGATGATG 3060
Db 1980 AGGTCAGTCAGAGCTGACGATTCAGAGCACTTCCAGAAATGATGATGATGATGATGATG 2039
Qy 3061 GAATATGATATCAAGTGCAGAGATTAATGATGATGATGATGATGATGATGATGATGAT 3120
Db 2040 GAATATGATATCAAGTGCAGAGATTAATGATGATGATGATGATGATGATGATGATGAT 2099
Qy 3121 AGAAATGTTTACAGAGTGAAGGAAATTAATTAATTAATTAATTAATTAATTAATTAAT 3180
Db 2100 AGAAATGTTTACAGAGTGAAGGAAATTAATTAATTAATTAATTAATTAATTAATTAAT 2113
Qy 3181 TGGAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 3240
Db 2114 ----- 2128
Qy 3241 GAATCTGAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3300
Db 2129 GAATCTGAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2188
Qy 3301 ATTATACACTGTTGAACTATCTGTTGACAGTGTGACAGTGTGACAGTGTGACAGTGTG 3360
Db 2189 ATTATACACTGTTGAACTATCTGTTGACAGTGTGACAGTGTGACAGTGTGACAGTGTG 2248
Qy 3361 CTCTGTGAGAGCACTGAGAGAGTGTGCTGTGATGATGATGATGATGATGATGATGATG 3420
Db 2249 CTCTGTGAGAGCACTGAGAGAGTGTGCTGTGATGATGATGATGATGATGATGATGATG 2308

Qy 3421 ACATGCTACTGATGTCAGAACTTACACAAAGATATGAGATTTTCTGCTCAAGTCAG 3480
Db 2309 ACATGCTACTGATGTCAGAACTTACACAAAGATATGAGATTTTCTGCTCAAGTCAG 2368
Qy 3481 GGTGAAATCACTGAGAGCTTGTCTTGCATCATTA 3518
Db 2369 GGTGAAATCACTGAGAGCTTGTCTTGCATCATTA 2406

RESULT 8
US-10-600-070-132/c
; Sequence 132, Application US/10600070
; Publication No. US20040139500A1
; GENERAL INFORMATION:
; APPLICANT: OsteYoung, Katherine W.
; APPLICANT: Vittha, Stanislaw
; APPLICANT: Kokeharova, Olga A.
; APPLICANT: Gao, Honggo
; TITLE OF INVENTION: Placid Division and Related Genes and Proteins, and Methods of
; FILE REFERENCE: MSU-08153
; CURRENT APPLICATION NUMBER: US/10/600,070
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 132
; LENGTH: 561
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (127)..(127)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (520)..(520)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (541)..(541)
; OTHER INFORMATION: n is a, c, g, or t
; US-10-600-070-132

Query Match 11.0%; Score 402.2; DB 7; Length 561;
Best Local Similarity 82.2%; Pred. No. 6.1e-95;
Matches 514; Conservative 0; Mismatches 20; Indels 91; Gaps 1;
Qy 2970 TTTCTTCTGATTTTCTTCTTCTTCTGATTTAGGCTCAGTCAGACGATTCAGAAAG 3029
Db 534 TTTCTTCTGATTTTCTTCTTCTTCTGATTTAGGCTCAGTCAGACGATTCAGAAAG 475
Qy 3029 CACTTCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3089
Db 474 CACTTCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 415
Qy 3090 AGTCTGAGCTTTTGGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3149
Db 414 AGTCTGAGCTTTTGGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 370
Qy 3150 TCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3209
Db 369 ----- 370
Qy 3210 TGATTTCTTATTTATATGATTTTGAATGAGGAAATGCTGAAGATTTTGAATGACATGAC 3269
Db 369 ----- 370
Qy 3270 GCTGAACTGCGAGCTGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 3329
Db 325 GCTGAACTGCGAGCTGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 266
Qy 3330 AGTGTGACATGTCAGACAGATGAACCCGTGCTGTGTGAAAGCACTTGTGAGAGATCT 3389


```
QY 1007 AATAAGTTCTTCTGTTTAATTTCAATGATGAATAAGAAAGAACTTTATCTAGTG 1066
DB 620 -----TCGAC 624
QY 1067 AAGGTTCTGGGGCTCTGTGTATGCAAGAGGTGTAGACATGATGTTCTGGG 1126
DB 625 AAGTTCTGGAGCCGTGTGTGTGTGTCAGAGAGCTGAGAGACGAGCTTGTGCTGAG 684
QY 1127 GTTGTGAGGCTCTGTGTTAAGAGAGGTGCTTAAGCTGTTTAAAGCAAGATGTGTTTAA 1186
DB 685 ATTGGGCAAGGTTGTGCTTAGGAGAGGTTGCCGAAGACGTTTAAAGAGATGTGTGTG 744
QY 1187 GTTATGAGCCCTGCTGTTTCTCATGTCTCGAGGAGATGCTATGSCATTGATCCACTGAT 1246
DB 745 GGTATGAGCACTGCAATTTGTATGAGTGTCAAGGAGATGCTATGSCCTTGTCTCACCGAGT 804
QY 1247 TTTATTAATCTGTTATGAGTTGTGTGAGAGAGCTTGAAGCTTTTACAGGATGTTGACTT 1306
DB 805 TTCAATTGCGGCTGTGAGATGCTCGAGAGGCGTTGAGCTTTTGCAAGAA----- 855
QY 1307 GCTTGTGTAATTGACGAGCGTTGCTTATTAAGAACTTTCTGATTTGATCTTTGTAA 1366
DB 856 ----- 855
QY 1367 TTGAGTCTTGTGTAGAGAGAGCAAGTACCTTGACCCGAGTTTACGTGCACAATT 1426
DB 856 -----GAGGGGCAACCACTGATGCTCCAGATTTCACACACAATA 897
QY 1427 GATGAGCTTTGGAAGAGATCACTCCGCGTTATGCTTGTGAGCTTCTGCGCTTACCGCTT 1486
DB 898 GATGAGACCTCTGAGAGATTAACCCACGTTGTGTGGAACCTTTTACCTTGCCCTCTT 957
QY 1487 GGTATGATTAATGCTGCGAAGAAAGACTAAATGCTTTAAGCGTGTGCGAATATTTTGTG 1546
DB 958 GATGACGAACATCGAGCGCGGAGGAGAGAGGCTCTTGTGTGTGCGTAACATTTTGTG 1017
QY 1547 TCTGTTGAGAGAGGTGTGAGCATCATGCTTGTGTGGGGTTTAAACCGGTGAGAGATTATG 1606
DB 1018 GCGGTGTGTGAGAGGGGTGTGCAACAATTTGCCGGGTTTACCCGTGAMAGACTTCAATG 1077
QY 1607 AATGAGGCGTTTTCAGATGACAGTGTGTGAGCGTTATACAGTTTGAATGCC 1660
DB 1078 AATGAGGCAATCTTTCACATGACAGCGCTGAACAGGTTGAACCTTTTGTGTACC 1131

RESULT 11
US-10-600-070-135
; Sequence 135, Application US/10600070
; Publication No. US20040139500A1
; GENERAL INFORMATION:
; APPLICANT: Osteoryoung, Katherine W.
; APPLICANT: Vitsha, Stanislaw
; APPLICANT: Kokeharova, Olga A.
; APPLICANT: Gao, Hongo
; TITLE OF INVENTION: Plasmid Division and Related Genes and Proteins, and Methods of
; FILE REFERENCE: MSU-08153
; CURRENT APPLICATION NUMBER: US/10/600,070
; CURRENT FILING DATE: 2003-06-20
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 135
; LENGTH: 660
; TYPE: DNA
; ORGANISM: Medicago truncatula
US-10-600-070-135

Query Match 5.7%; Score 208.8; DB 7; Length 660;
Best Local Similarity 60.3%; Pred. No. 7.8e-44;
Matches 429; Conservative 0; Mismatches 207; Indels 76; Gaps 2;
561 AAGTAGCCACAACACTCTATCACTATCTGCTCCGCGCAAAATGGCGGACCGTCTTCT 620
|||||
```

```
DB 17 ACCTAACCGTCTCCATTTCTCCGCGGTCTCCGCCACAGTAATGGCGGAGCACTCAT 76
QY 621 CTCCGACTTCAATTTCACTCCGATTTCCCTCTCTCTCTCTCTGCGACCGCACCAAC 680
DB 77 TTCCGATTTTCAATTTCTCTGAGACACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 136
QY 681 CGCCACTCTGCTCTCTCTGCGACATCTATGTATGTATGTATGTATGTATGTATGTATGTAT 740
DB 137 AGTCACCTCTC-----ACTCTCTCTTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 190
QY 741 CATGATTTTCAACAGATTAATGAGTCAACCAACATTTCTTAAACCGTATGATGAAGAG 800
DB 191 TCTGACCTGTACAAATCTCTGCGCGCAAAACGATTTTCTCGGTATGATGTATGAG 250
QY 801 AGCATTCGAAGTATGAGGTTTGTGAAACCGCCCAATTCGGTTTCAAGCGACGACTTTAAT 860
DB 251 AGCTTATGAGAGAAATCTCTGAAACCTCTCAGATATCTTCAATATGAAAGCTTTGAT 310
QY 861 CAGCCGAGACAGATTTTCAAGCTGTGCGAAACCTGTCTAATTTCTCGGCTTGAAG 920
DB 311 TAGTGTGTGAGATTTCTTCAAGCTGTGTGAAACCTTACGTATCTGCTTCTAGAAG 370
QY 921 AGAGTACAAATGAGTCTTCTTGTATGATGAAGAAAGCTACATCATGATGTTCTTCT 980
DB 371 AGAGTATTAATCAAAAGCTCTGTGACAGATGAAGAGC----- 405
QY 981 GGATTAAGGTAATTTGATTTGCGAATATAATTAAGTTTCTTGTATTAATTCATGAATTG 1040
DB 406 -----AAGATGAGAAATCTT 420
QY 1041 AATAAGAGAACTTTATCTATGTAAGTTCCTGCGGCTCTCTGTGTATTTCAAGAG 1100
DB 421 CCATTTCTACGAAATCCCTTTTCGCAAAAGTTCCGAGACTCTGTGCGTGTGCAAGAG 480
QY 1101 GTGTGAGACATGAGATTAATTTCTGCGGTGTGTGAGGCTCTGTATGAAGAGGTTGCC 1160
DB 481 CTGAGAGACGAGATGTGTGTCTGAGATTTGAGAGGGGTTTACTGAGAGAGAGTTACCGA 540
QY 1161 AGTCGTTTAAACAAGATGTGTTTATGTTATGTATGAGCGCTTGTCTCATGTCTCGAGGG 1220
DB 541 AGATGTTTAAACAGATTTGTGTGTGCTATGCGCTTGTGCTATGCTTGAATTTCTAAGGG 600
QY 1221 ATGCTANGCAATTGATTCACCTGATTTTATTAATTAATGTTATGATTTGTGA 1272
DB 601 ATGCTATGCTTTGTGTCCCGCCAGATTTCAATGTGTTGTGTGAGATGTCTGA 652

RESULT 12
US-10-600-070-126
; Sequence 126, Application US/10600070
; Publication No. US20040139500A1
; GENERAL INFORMATION:
; APPLICANT: Osteoryoung, Katherine W.
; APPLICANT: Vitsha, Stanislaw
; APPLICANT: Kokeharova, Olga A.
; APPLICANT: Gao, Hongo
; TITLE OF INVENTION: Plasmid Division and Related Genes and Proteins, and Methods of
; FILE REFERENCE: MSU-08153
; CURRENT APPLICATION NUMBER: US/10/600,070
; CURRENT FILING DATE: 2003-06-20
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 126
; LENGTH: 2283
; TYPE: DNA
; ORGANISM: Oryza sativa
US-10-600-070-126

Query Match 5.3%; Score 194; DB 7; Length 2283;
Best Local Similarity 55.6%; Pred. No. 1.4e-39;
Matches 514; Conservative 0; Mismatches 300; Indels 110; Gaps 3;
```


OTHER INFORMATION: Clone ID: LIB3493-032-PI-M1-F8
US-09-732-627A-2154

Query Match 4.8%; Score 175; DB 3; Length 439;
Best Local Similarity 65.0%; Pred. No. 5e-35;
Matches 278; Conservative 0; Mismatches 141; Indels 9; Gaps 1;

```
QY 499 GTCCGATGTCCTCTCCCAATTCATTAATGCGGATTAACACCGCGGACGACAAAGCTC 558
DB 12 GTTAGATCGGGCTGTCACCTCCGACACTGCTCTTATGCAACCTGTAAACCTCTC 71
QY 559 CGACGTAGCCACACACCTCTCAACTATCTGTCGCGACGAATGGGCGGACCGCTTT 618
DB 72 ANACTCCACACCGTTCAACTACCGTCTGCTCGGCAATGAATGGGCGGACCGCTTC 131
QY 619 CTCTCCGATTCATTAATTCACCTCG-----ATTCTCTCTCTCTCTCTCTCTCTCTC 669
DB 132 CTGTGCTGACTTCATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTA 191
QY 670 GCGACCAACCAACCGCACTCTCTCTCTCTGCGACCACTATTGATGTCGCGACCGCAC 729
DB 192 ACCGCAACTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 251
QY 730 GTCCCATCTCCATGATTTCTTACAGGATTAAGAGCTCAACACATTTCTTAACCAT 789
DB 252 GTTTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 311
QY 790 GGAATCAGAAAGACATTCGAACTAGGGTTTGAACCGCGCAATTCGGTTTCAAGCAG 849
DB 312 GGAATTAAGAGAGCTTATTAAGAGAGGTTTGAACCGCGCTCATATAGGTTCAAGTCA 371
QY 850 GACGCTTTAATCAGCGCGGAGACAGATTTCTTAAGCTCTTGGGAAACTCTGTATCT 909
DB 372 GACACCATTAATTAAGCGGAGACAGATTTCTTAAGCTCTGCTGGAACCTATCTAACCT 431
QY 910 CGGTCTAG 917
DB 432 GGGCTCTAG 439
```

RESULT 15
US-10-600-070-175

Sequence 175, Application US/10600070
Publication No. US20040139500A1
GENERAL INFORMATION:
APPLICANT: Oeteryoung, Katherine W.
APPLICANT: Vitka, Stanislaw
APPLICANT: Kokeharova, Olga A.
APPLICANT: Gao, Hong
TITLE OF INVENTION: Placid Division and Related Genes and Proteins, and Methods of
FILE REFERENCE: MSU-08153
CURRENT APPLICATION NUMBER: US/10/600,070
CURRENT FILING DATE: 2003-06-20
NUMBER OF SEQ ID NOS: 206
SOFTWARE: PatentIn version 3.2
SEQ ID NO 175
LENGTH: 545
TYPE: DNA
ORGANISM: Gossypium arboreum
FEATURE:
NAME/KEY: misc_feature
LOCATION: (528)..(528)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (536)..(536)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (540)..(540)
OTHER INFORMATION: n is a, c, g, or t
US-10-600-070-175

Query Match 4.4%; Score 161; DB 7; Length 545;
Best Local Similarity 69.6%; Pred. No. 2.9e-31;
Matches 218; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

```
QY 3224 ATAGTTTGGATGGGCGAATGCTGAAGATTTGGACTGACAGACGACTGAAACTGGCGA 3283
DB 82 AGAGGTTCTGATGATGTCAAATGTTGAAAGACATGACATGCTGCGAAGCGAAATGCTCA 141
QY 3284 GCTTGGTTGTTATGATTAATTAACCTGTTGAAACTATCTGTTGACAGTGAAGTCTC 3343
DB 142 GCTTGGTTGTTATGATTAATTAAGTCTGTTGAAACTATCTGTTGACAGTGAAGTCTC 201
QY 3344 AGCAGATGAAACCGCTGCTCTGTTGGAACAACCTGAGAGAGTCTGTTGTTATCTGA 3403
DB 202 ACTAGATGGCCAGGAGGCTGATGTAAGTACTGTAAGTCTGTAAGATTCACCTGTTGACTGA 261
QY 3404 TTTGTTTATCGAAGAAACATGCTAGTGTGCAACCTGACACCAAGATGCAAGT 3463
DB 262 TGTTCATCATCGGAGAACATGCTCTAATGTAACTCTTAACACAGAGATGAGAT 321
QY 3464 TTTCTGTCGAAGTCAGGAGTGAATAATCACTGAAGGCTCTGTTTGCATCATATATATAC 3523
DB 322 GTCTTGTTCGAACCTGAGGCTGGAATATCACTGAGATCTGTCAAAATCTTAATCTATG 381
QY 3524 TCATATGATGAT 3536
DB 382 ATGTATTAAGCAT 394
```

Search completed: December 10, 2005, 18:16:03
Job time : 2829.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 10, 2005, 03:43:47 ; Search time 611.67 Seconds
(without alignments)
10656.596 Million cell updates/sec

Title: US-10-600-070b-3

Perfect score: 3667
Sequence: 1 Tgtctgcatgaagagaat.....ctatacataagagctacaa 3667

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1303057 segs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: /cgn2_6/prodata/1/ina/1_COMB.seq:*
2: /cgn2_6/prodata/1/ina/5_COMB.seq:*
3: /cgn2_6/prodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/prodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/prodata/1/ina/H_COMB.seq:*
6: /cgn2_6/prodata/1/ina/PCUS_COMB.seq:*
7: /cgn2_6/prodata/1/ina/PP_COMB.seq:*
8: /cgn2_6/prodata/1/ina/RB_COMB.seq:*
9: /cgn2_6/prodata/1/ina/Backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	57.6	1.6	7218	2	US-08-232-463-14
2	54	1.5	1141	3	US-09-806-708B-22
3	50.2	1.4	57280	3	US-09-949-016-11796
4	50.2	1.4	57280	3	US-09-949-016-11796
5	50.2	1.4	57280	3	US-09-949-016-12843
6	50.2	1.4	57280	3	US-09-949-016-12843
7	50.2	1.4	57280	3	US-09-949-016-12844
8	50.2	1.4	57280	3	US-09-949-016-13542
9	50.2	1.4	57280	3	US-09-949-016-13543
10	50.2	1.4	57280	3	US-09-949-016-13544
11	50.2	1.4	57280	3	US-09-949-016-13545
12	50.2	1.4	57280	3	US-09-949-016-14633
13	50.2	1.4	57280	3	US-09-949-016-14634
14	50.2	1.4	57280	3	US-09-949-016-14635
15	50.2	1.4	57280	3	US-09-949-016-14636
16	50.2	1.4	57280	3	US-09-949-016-14637
17	50.2	1.4	57280	3	US-09-949-016-14638
18	50.2	1.4	57280	3	US-09-949-016-14639
19	49.4	1.3	913	2	US-08-217-327-3
20	49.4	1.3	913	2	US-07-885-970A-3
21	49.4	1.3	913	2	US-08-298-687A-3
22	49.4	1.3	913	2	US-08-530-797-2
23	49.4	1.3	913	2	US-08-298-829-3
24	49.4	1.3	913	2	US-08-787-335-2

25	49.4	1.3	1984	2	US-07-885-970A-25	Sequence 25, Appl
26	49.4	1.3	1985	2	US-08-298-687A-25	Sequence 25, Appl
27	49.4	1.3	1985	2	US-08-298-829-25	Sequence 25, Appl
28	49.2	1.3	324	3	US-09-547-693-234	Sequence 234, Appl
29	48.4	1.3	152132	3	US-09-949-016-13845	Sequence 12371, A
30	48.4	1.3	152145	3	US-09-949-016-13371	Sequence 12371, A
31	47.4	1.3	1141	3	US-09-806-708B-22	Sequence 22, Appl
32	46.8	1.3	114793	3	US-10-148-806-3	Sequence 12458, A
33	45	1.2	49818	3	US-09-949-016-1458	Sequence 14011, A
34	45	1.2	49829	3	US-09-949-016-14011	Sequence 2, Appl
35	44.2	1.2	4403765	3	US-09-103-840A-2	Sequence 2813, Ap
36	44.2	1.2	441529	3	US-09-103-840A-1	Sequence 2813, Ap
37	43.8	1.2	832	3	US-09-621-976-2813	Sequence 23, Appl
38	43.8	1.2	832	3	US-09-621-976-2813	Sequence 14080, A
39	42.8	1.2	1055	3	US-09-806-708B-23	Sequence 13813, A
40	42.8	1.2	60990	3	US-09-949-016-14080	Sequence 12384, A
41	42.4	1.2	23849	3	US-09-949-016-13813	Sequence 12801, A
42	42.4	1.2	39154	3	US-09-949-016-12384	Sequence 14326, A
43	42.4	1.2	39154	3	US-09-949-016-12801	Sequence 14327, A
44	42.4	1.2	39443	3	US-09-949-016-14326	
45	42.4	1.2	39443	3	US-09-949-016-14327	

ALIGNMENTS

RESULT 1
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOULPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESS: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-9300
; TELEFAX: (703) 683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZspt-Fls

Best Local Similarity 53.9%; Pred. No. 0.033;
Matches 103; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY	565	AGGACACAAACCCCTTCAATATATCTGGTCCGACAGAAATGGGCGACCGTCTTCTCC	624
Db	13101	AGGACCATACACACACAACTGGCTCCACCAATACCAACCTTCAACATTATAGACTCC	13042
QY	625	GACTTCATTTCACCTTCGATTCCTCTCTCTCTTTCGACACGACACCAACCGCC	684
Db	13041	ACCATCAATATCCCTATTCAGCTCCACCAACCGCTTTCACATACACACACACACAC	12982
QY	685	ACTCTGCTCTCTCTGGCACCATTTATTTGATGTCGCCGAAGCGACGCTCCCATCCCA	744
Db	12981	ACAACACACACTCATATATCATGACTTCCACCATCAACATTACCATCACTCCACCAAC	12922
QY	745	GATTTCACCA	755
Db	12921	CATCATCACGA	12911

RESULT 4
US-09-949-016-12843/c

```

Sequence 12843, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
PRIORITY FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12843
LENGTH: 57280
TYPE: DNA
ORGANISM: Human
US-09-949-016-12843

```

Query Match	1.4%;	Score 50.2;	DB 3;	Length 57280;
Best Local Similarity	53.9%;	Pred. No. 0.033;		
Matches 103; Conservative	0;	Mismatches 88;	Indels 0;	Gaps 0;

QY	565	AGGCACAAACACCTCTCAACTATATCTGTGCGCAGGAAATGGGCGACCGTCTCTCTCC	624
Db	13101	AGACACATACCAACCAACACTGCGCTTCACACACATACAACCTTTCACCATTAATAGCTCC	13042
QY	625	GACCTTCATTTTCACTTCGATTCCTCTCTCTCTCTCTGCGCACCGCACACACCGCC	684
Db	13041	ACCATCAATATCTCTTAACGTCAACACACACACCGCTTTACCACTACCAACACACACAC	12982
QY	685	ACTCTGCTCTCTGCAACCATCTAATGATGCTGCCGAAGCGACGTCCCATTCGCATT	744
Db	12981	ACAACACACACATCATATCATGACTTCCACACATCAATACATCATCCTCCACGCACAC	12922
QY	745	GATTTCTACCA	755
Db	12921	CATCATACCGA	12911

RESULT 5
US-09-949-016-12844/C

Sequence 12844, Application US/09945016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTNER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

```

? TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
? FILE REFERENCE: CLO01307
? CURRENT APPLICATION NUMBER: US/09/949,016
? CURRENT FILING DATE: 2000-04-14
? PRIOR APPLICATION NUMBER: 60/241,755
? PRIOR FILING DATE: 2000-10-20
? PRIOR APPLICATION NUMBER: 60/237,768
? PRIOR FILING DATE: 2000-10-03
? PRIOR APPLICATION NUMBER: 60/231,498
? PRIOR FILING DATE: 2000-09-08
? NUMBER OF SEQ ID NOS: 207012
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 12844
? LENGTH: 57280
? TYPE: DNA
? ORGANISM: Human
US-09-949-016-12844

Query Match          1.4%  Score 50.2;  DB 3;  Length 57280;
Best local Similarity 53.9%;  Pred. No. 0.033;
Matches 103;  Conservative 0;  Mismatches 88;  Indels 0;  Gaps 0;

```

Query Match	1.4%;	Score 50.2;	DB 3;	Length 57280;
Best Local Similarity	53.9%;	Pred. No. 0.033;		
Matches 103;	Conservative	0;	Mismatches 88;	Indels 0;
				Gaps 0;

QY	565	AGCGCAACAACCTCTACAACTCTATTCGCTCCGCGACGAAATGGCGGACGGTCTTCTCC	624
QY	13101	AGCAGCATATCACACCAAACTGCTCTCCACACCATACCACTTACCATTTATGAGCTCC	13042
Db	625	GACTTCAAATTTGACCTCGATTCTCTCTCTCTCTCTCTCTGCGACCGCACCAACCGCC	684
QY	13041	ACCATCAATATCTCCATATCAAGTCCGACGACCAACCGTCTTACCATACGACGACGACAC	12982
Db	685	ACTCTCGTCTCTCTGCGACCATCTATTGATGTCGCCGAAGCGACGTCCTCCATTC	744
QY	12981	ACAACACCACTCACTATCATCATGATTCACACATCAACATTACATCACTCCACCAACAAC	12922
Db	745	GATTTCTACCA	755
QY	12921	CATCATCAACGA	12911
Db			

RESULT 6

```

US-09-949-016-12846/c / Sequence 12846, Application US/09949016
/ Patent No. 6812319
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ CURRENT FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 12846
/ LENGTH: 57280
/ TYPE: DNA
/ ORGANISM: Human
/ US-09-949-016-12846

```

Query Match	1.4%	Score 50.2;	DB 3;	Length 57280;
Best Local Similarity	53.9%;	Pred. No. 0.033;		
Matches 103; Conservative	0;	Mismatches 88;	Indels 0;	Gaps 0;

Qy 565 AACCAACAACTCTACAACTATCTGCTCTCGCAGAAATGGGCGACCGCTCTTCTCTCC 624

Db 13101 AACCACTATCCACCAACAACTGCTCTCCACCAACATACCACTTACCACTTATGAGCTCC 13042

Qy 625 GACTTCAAATTTGACCTCGATTCTCTCTCTCTCTTCGCGACCGCGACCAACACCGCC 684

Db 13041 ACCATCAATATCCCTATACAGTCCACCAACCCCTCTTCCACCATCACCACCAACC 12982
QY 685 ACTCTGCTCTCTGCGACCATCTATGATGTCGCGAAGCCGACGTCCTTCTCTCC 744
Db 12981 ACAACCAACCACTCACTATCATGACTTCCACCATTAACATTCATCCTCCACCAAC 12922
QY 745 GATTTCTACCA 755
Db 12921 CATCATCAGCA 12911

RESULT 7
US-09-949-016-13542/c
; Sequence 13542, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 13542
; LENGTH: 57280
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13542

Query Match 1.4%; Score 50.2; DB 3; Length 57280;
Best Local Similarity 53.9%; Pred. No. 0.033;
Matches 103; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 565 AGCCACAACACCTTCACTATCTGCTCCGCGCAAGTGGCCGACCGTCTTCTCTCC 624
Db 13101 AGCACCAATCAACCAACAGCTCCACCACTCAACCACTTACCATATGAGCTCC 13042
QY 625 GACTCAATTTCACTCCGATTCCTCTCCCTCCCTTGGCCAGCGACCAACCGCC 684
Db 13041 ACCATCAATATCCCTATACAGTCCACCAACCGCTTTCACCATCACCACCAAC 12982
QY 685 ACTCTGCTCTCTGCGACCATCTATGATGTCGCGAAGCCGACGTCCTTCTCTCC 744
Db 12981 ACAACCAACCACTCACTATCATGACTTCCACCATTAACATTCATCCTCCACCAAC 12922
QY 745 GATTTCTACCA 755
Db 12921 CATCATCAGCA 12911

RESULT 8
US-09-949-016-13543/c
; Sequence 13543, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 13543
; LENGTH: 57280
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13543

Query Match 1.4%; Score 50.2; DB 3; Length 57280;
Best Local Similarity 53.9%; Pred. No. 0.033;
Matches 103; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 565 AGCCACAACACCTTCACTATCTGCTCCGCGCAAGTGGCCGACCGTCTTCTCTCC 624
Db 13101 AGCACCAATCAACCAACAGCTCCACCACTCAACCACTTACCATATGAGCTCC 13042
QY 625 GACTCAATTTCACTCCGATTCCTCTCCCTCCCTTGGCCAGCGACCAACCGCC 684
Db 13041 ACCATCAATATCCCTATACAGTCCACCAACCGCTTTCACCATCACCACCAAC 12982
QY 685 ACTCTGCTCTCTGCGACCATCTATGATGTCGCGAAGCCGACGTCCTTCTCTCC 744
Db 12981 ACAACCAACCACTCACTATCATGACTTCCACCATTAACATTCATCCTCCACCAAC 12922
QY 745 GATTTCTACCA 755
Db 12921 CATCATCAGCA 12911

RESULT 9
US-09-949-016-13544/c
; Sequence 13544, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 13544
; LENGTH: 57280
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13544

Query Match 1.4%; Score 50.2; DB 3; Length 57280;
Best Local Similarity 53.9%; Pred. No. 0.033;
Matches 103; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 565 AGCCACAACACCTTCACTATCTGCTCCGCGCAAGTGGCCGACCGTCTTCTCTCC 624
Db 13101 AGCACCAATCAACCAACAGCTCCACCACTCAACCACTTACCATATGAGCTCC 13042
QY 625 GACTCAATTTCACTCCGATTCCTCTCCCTCCCTTGGCCAGCGACCAACCGCC 684
Db 13041 ACCATCAATATCCCTATACAGTCCACCAACCGCTTTCACCATCACCACCAAC 12982
QY 685 ACTCTGCTCTCTGCGACCATCTATGATGTCGCGAAGCCGACGTCCTTCTCTCC 744
Db 12981 ACAACCAACCACTCACTATCATGACTTCCACCATTAACATTCATCCTCCACCAAC 12922
QY 745 GATTTCTACCA 755


```

? GENERAL INFORMATION:
? APPLICANT: VENTER, J., Craig et al.
? TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
? TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
? FILE REFERENCE: C1001307
? CURRENT APPLICATION NUMBER: US/09/949,016
? CURRENT FILING DATE: 2000-04-14
? PRIOR APPLICATION NUMBER: 60/241,755
? PRIOR FILING DATE: 2000-10-20
? PRIOR APPLICATION NUMBER: 60/237,768
? PRIOR FILING DATE: 2000-10-03
? PRIOR APPLICATION NUMBER: 60/231,498
? PRIOR FILING DATE: 2000-09-08
? NUMBER OF SEQ. ID NOS: 207012
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ. ID NO 14635
? LENGTH: 57280
? TYPE: DNA
? ORGANISM: Human
? OS-09-949-016-14635

```

Query Match	1.4%	Score 50.2	DB 3	Length 57280
Best Local Similarity	53.9%	Pred. No. 0	0.033	
Matches 103	Conservative 0	Mismatches 88	Indels 0	Gaps 0
QY	565	AGCCACAACCTCTTACACTATCTGCTCCGACGCAAAATGGCGCAGCGCTCTCTCC	624	
Db	13101	AGCACCAATACACACCAACACTGCTCTCCACCAACATCACCACTTACCAATTATGAGCTCC	13044	
QY	625	GACATCAATTTTCACCTCCGATTCCTCTCTCTCTGGACGCGGCACACACAGCGCC	684	
Db	13041	ACCATTCAATATCTCTTATCAAGTTCACACCAACCGTCTTACACATACACACACACACC	12982	
QY	685	ACTCTCGCTCTCTGACCAACATCTATGTATCGTCCGAAAGCGACGTCGCCATCCCATTT	744	
Db	12981	ACAACCAACACTCACTATCAAGCTTCCACCATCAACATTATACCTACCGTCGACGACAC	12922	
QY	745	GATTTTCAACCA	755	
Db	12921	CATCATCAACGA	12911	

```

RESULT 14
US-09-949-016-14636/C
; Sequence 14636, Application US/09949016
; Patent No. 6812335
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C10010307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14636
; LENGTH: 57280
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-14636

```

Query Match	1.4%	Score 50.2	DB 3	Length 57280
Best Local Similarity	53.9%	Pred. No. 0.033		
Matches 103	Conservative	0	Mismatches 88	Indels 0
Gaps	0			
565	AGCCACACACTCTACACACTATCTGCTCCGCGACGAAATGGGGCGACCGCTTTCCTCC	624		

Db	13101	AGACATCAACGACGACAACTGCTCCACACACATACACACTTACCATTTAGAGCTCC	13042
Qy	625	GACTTCAATTTACCTTCGATTCTCTCTCTCTTGGCCACGGCCACGACCGCC	684
Db	13041	ACCATCATATTCCTTATCGATCGACACACACCGTTCATCCATCAACGACGACGCC	12982
Qy	685	ACCTCTGCTCTCTGGCACCATATTGATGCTGCCAGCGACGCTCCCATCCCAT	744
Db	12981	ACAACCAACCACTCATATCATGACTTCCACCATCAACATTACCATCACTCCACCAAC	12922
Qy	745	GATTTCTACCA	755
Db	12921	CATCATCACA	12911

```

RESULT 15
US-09-949-016-14637/C
? Sequence 14637, Application US/09949016
? Patent No. 681339
? GENERAL INFORMATION:
? APPLICANT: VENTER, J. Craig et al.
? TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
? TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
? FILE REFERENCE: C1001307
? CURRENT APPLICATION NUMBER: US/09/949,016
? CURRENT FILING DATE: 2000-04-14
? PRIOR APPLICATION NUMBER: 60/241,755
? PRIOR FILING DATE: 2000-10-20
? PRIOR APPLICATION NUMBER: 60/237,768
? PRIOR FILING DATE: 2000-10-03
? PRIOR APPLICATION NUMBER: 60/231,498
? PRIOR FILING DATE: 2000-09-08
? NUMBER OF SEQ ID NOS: 207012
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 14637
? LENGTH: 57280
? TYPE: DNA
? ORGANISM: Human
? US-09-949-016-14637

```

[illegible]

Search completed: December 10, 2005, 08:09:24
Job time : 617.67 secs

Search completed: December 10, 2005, 08:09:24
Job time : 617.67 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 10, 2005, 16:58:05 ; Search time 13948.2 Seconds
(without alignments)
12300.344 Million cell updates/sec

Title: US-10-600-070b-3

Perfect score: 3667
Sequence: 1 tgcctgcataaagagagaat.....ctataacataagggctacaa 3667

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

BST: *
1: gb_est1: *
2: gb_est2: *
3: gb_est3: *
4: gb_hic: *
5: gb_est4: *
6: gb_est5: *
7: gb_est6: *
8: gb_est7: *
9: gb_gsa1: *
10: gb_gsa2: *
11: gb_gsa3: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length	ID	Description
1	725.2	19.8	919	4	CNS09YJH
2	436.2	11.9	741	4	CNS09Z5H
3	434.6	11.9	741	4	CNS09YJH
4	402.2	11.0	561	1	AI998415
5	376.4	10.3	415	3	BP782886
6	363.6	9.9	785	3	BP785511
7	359.2	9.8	434	3	BP620404
8	359.2	9.8	434	3	BP620404
9	359.2	9.8	434	3	BP620404
10	359.2	9.8	434	3	BP620404
11	334	9.1	871	10	DU050213
12	315.8	8.6	534	1	AV830764
13	311.8	8.5	320	9	BZ765587
14	307.2	8.4	320	9	BZ765587
15	305.6	8.3	320	9	BZ765587
16	304.4	8.3	320	9	BZ765587
17	299.4	8.2	311	8	AY199896
18	295.6	8.1	631	5	BU046755
19	293.4	8.0	832	7	CO079829
20	265.4	7.2	751	11	CR486014
21	254	6.9	849	7	CO117047
22	252.6	6.9	769	7	CO075595

23	252	6.9	813	7	CO071968
24	244.4	6.7	741	8	CX543653
25	239.6	6.5	423	1	AV812946
26	237.2	6.5	703	7	CN918842
27	226	6.2	728	8	CX674370
28	212.6	5.8	521	7	CN914655
29	212.6	5.8	593	7	CN904734
30	210	5.7	215	10	AL952467
31	208.8	5.7	660	2	BI268376
32	202.8	5.5	697	8	CX173275
33	194	5.3	2307	10	CL965374
34	193.2	5.3	510	6	CF603268
35	166.4	4.5	542	5	BQ586205
36	161.8	4.4	819	7	CV241483
37	161.8	4.4	821	7	CK090561
38	161.8	4.4	922	7	CV264112
39	161	4.4	545	5	BQ410206
40	161	4.4	692	10	CL723297
41	158.8	4.3	831	8	DN939692
42	158	4.3	608	2	BE472035
43	157.8	4.3	722	7	CO117046
44	157.4	4.3	321	10	AL936312
45	156.2	4.3	947	10	CG088802

ALIGNMENTS

RESULT 1
CNS09YJH 919 bp mRNA linear HTC 04-FEB-2004
LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone
DEFINITION GSLT89ZC08 of Adult vegetative tissue of strain col-0 of
Arabidopsis thaliana (thale cress).

ACCESSION BX841670.1 GI:42406830
VERSION BX841670.1
KEYWORDS HTC; GSLT cDNA.
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana (thale cress)

REFERENCE
AUTHORS Castel, V., Aury, J.M., Jallion, O., Wincker, P., Clepet, C., Menard, M., Cruaud, C., Quelet, F., Searpelli, C., Schachet, V., Temple, G., Caboche, M., Weissenbach, J. and Salanoubat M.
Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation

TITLE
JOURNAL
REFERENCE
AUTHORS
JOURNAL
COMMENT

Unpublished
2 (bases 1 to 919)
Direct Submission
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castel, V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C., Schachet V., Weissenbach J., Salanoubat M.
URCV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Pnap.
http://www.genoscope.cns.fr/externe/sequences/Banque_projet_EF/Full
length

http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.
location/qualifiers
1. 919
/organism="Arabidopsis thaliana"
/mol_type="mRNA"

FEATURES
source

/db_xref="taxon:3702"
/clone="GSLTIL31ZC08"
/tissue_type="Adult vegetative tissue"
/ecotype="Col-0"
/plasmid="pCMVSPORT_6"
1. 919
/gene="At5g42480"

ORIGIN

Query Match 19.8%; Score 725.2; DB 4; Length 919;
Best Local Similarity 84.2%; Pred. No. 7e-176;
Matches 916; Conservative 0; Mismatches 3; Indels 169; Gaps 2;

469 AAAGACATCCCATGGAAGCTCTGAGTCACTGGCATTTGTTCTCCCATTCATTA 528
1 AAAGACATCCCATGGAAGCTCTGAGTCACTGGCATTTGTTCTCCCATTCATTA 60
529 TGGCGATTACACCGCGGACGACAAAGTCCGACCTACCCCAACACTTCAACTTC 588
61 TGGCGATTACACCGCGGACGACAAAGTCCGACCTACCCCAACACTTCAACTTC 120
589 TGCTCCGACGAAATGGGCGGACCGCTCTCTCCGACTTCAATTCACTCCGATTTC 648
121 TGCTCCGACGAAATGGGCGGACCGCTCTCTCCGACTTCAATTCACTCCGATTTC 180
649 TCTCTCTCTCTCTGCGACCGGACGACACACCGGACCTCTCTCTCTGCGACCATCT 708
181 TCTCTCTCTCTCTGCGACCGGACGACACACCGGACCTCTCTCTCTGCGACCATCT 240
709 ATTGATGTCGCGAAGCGGACGCTCCCATTCCTCATTTCTCAAGATTTAGGAGCT 768
241 ATTGATGTCGCGAAGCGGACGCTCCCATTCCTCATTTCTCAAGATTTAGGAGCT 300
769 CAACACATTTCTTAACCGATGATCAAGAGAGCATTCGAAGCTAGGGTTTCGAACCG 828
301 CAACACATTTCTTAACCGATGATCAAGAGAGCATTCGAAGCTAGGGTTTCGAACCG 360
829 CCGCAATTCGTTTACGAGGACGCTTTAATCAGCCGGAACAGATTCTTCAAGCTCT 888
361 CCGCAATTCGTTTACGAGGACGCTTTAATCAGCCGGAACAGATTCTTCAAGCTCT 420
889 TGGCGAACTCTGCTAATCTCTGCTAAGAGAGTCAATGAAGCTCTTCTGATGAT 948
421 TGGCGAACTCTGCTAATCTCTGCTAAGAGAGTCAATGAAGCTCTTCTGATGAT 480
949 GAAGAAGCTACAGTCACTGATGATGCTTGGGATTAAGTAAATTCGATTTCGGAATTA 1008
481 GAAGAAGCTACAGTCACTGATGATGCTTGGGATTAAGTAAATTCGATTTCGGAATTA 516
1009 TAAAGTTCTCTGTTTAAATTCATGAATGGATTAAGGAAGAACTTTATCTAGTGA 1068
517 -----AA 518
1069 GATTCTGAGGCTCTCTGATATTGCAAGAGGTGAGAGCTGAGATGATCTTCCGGT 1128
519 GATTCTGAGGCTCTCTGATATTGCAAGAGGTGAGAGCTGAGATGATCTTCCGGT 578
1129 TGGTGAAGCTCTGCTTAAAGAGAGGTCTTAAGTCTTTAAGCAAGATGCTTTAGT 1188
579 TGGTGAAGCTCTGCTTAAAGAGAGGTCTTAAGTCTTTAAGCAAGATGCTTTAGT 638
1189 TATGCGCTTGGCTTCTGATGCTGAGAGGATCTTATGGCATTTGATCACTGATTT 1248
639 TATGCGCTTGGCTTCTGATGCTGAGAGGATCTTATGGCATTTGATCACTGATTT 698
1249 TATTAAGTGTATGATGCTTGTGAGAGAGCTTTGAAGCTTTTACAGAGTGAATTGCTTC 1308
699 TATTAAGTGTATGATGCTTGTGAGAGAGCTTTGAAGCTTTTAC----- 742
1309 TTTGTAATTTGACGAGGCTTGCTTTAAGAACTTTCTGATTTGATCTTTGTTATT 1368
743 ----- 742

1369 GAGTCTTGTTAGAGAGAGCAAGTACCTTGACCCGGATTTACGTGCACAAATTGA 1428
743 -----AGAGAGAGAGAGCAAGTACCTTGACCCGGATTTACGTGCACAAATTGA 791
1429 TGAGACTTTGGAAGAGATCACTCCGCTTATGCTTGGAGCTACTTGGCTTACCGCTTG 1488
792 TGAGACTTTGGAAGAGATCACTCCGCTTATGCTTGGAGCTACTTGGCTTACCGCTTG 851
1489 TGATGATTACGCTCGAAAAGACTAAATGGTTTAAAGCGGTGCGGAATATTTTGTGTC 1548
852 TGATGATTACGCTCGAAAAGACTAAATGGTTTAAAGCGGTGCGGAATATTTTGTGTC 911
1549 TGTGGAG 1556
912 TGTGGAG 919

RESULT 2
CNS09255
LOCUS
DEFINITION
Arabidopsis thaliana Full-length cDNA Complete sequence from clone
GSLTIL31ZB09 of Silique of strain col-0 of Arabidopsis thaliana
(thale cress).

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
BX833051.1 GI:42455503
HTC; GSLT_cDNA.
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids I; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE
AUTHORS
TITLE
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
Unpublished
2 (bases 1 to 741)
Genoscope.

JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqret@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castellani
V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences). 5 prime and 3 prime are assembled with phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
_length
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.
Location/Qualifiers
1. 741
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="GSLTIL31ZB09"
/tissue_type="Silique"
/ecotype="Col-0"
/plasmid="pCMVSPORT_6"
1. 741
/gene="At5g42480"

ORIGIN

Query Match 11.9%; Score 436.2; DB 4; Length 741;
Best Local Similarity 83.4%; Pred. No. 4.1e-101;

[illegible]

COMMENT

BP 191 91006 EURY cedex - FRANCE [E-mail : segrife@genoscope.cns.fr]
- Web : www.genoscope.cns.fr

The sequences are based on single pass reads.

Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G.

Genoscope members carried out sequencing and annotation : Castellani V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salmoubat M.

URCV INRA : Clepet C., Caboche M.

Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap.

http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Fulllength

<http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis>.

FEATURES

```

source
1. .741
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="GSLT51592A05"
/tissue_type="Silique"
/ecotype="Col-0"
/plasmid="pCMVSPORT_6"
1. .741
/gene="At5g42480"

```

ORIGIN

Query Match	11.9%	Score 434.6	DB 4	Length 741
Best Local Similarity	83.3%	Pred. No. 1.1e-100		
Matches 547; Conservative	0	Mismatches 19	Indels 91	Gaps 1

QY	2970	TTTTCTTCCCTGATTTTTTTTCTTTTGATTTAGGGTCAGTCAGACGCTGACGATTTGCAAG	3029
Db	176	TTTTCTTCTATGGAATCTGATGTCGCTACCATAGGGTCAGTCAGACGCTGACGATTTGCAAG	235
QY	3030	CACCTTCCAGAAATGATTCGTAGGACCTGCAGAGAAATATAGTATCAAGTGCAGAGATT	3089
Db	236	CATTTCCAGAAATGATTCGTAGGACCTGCAGAGAAATATAGTATCAAGTGCAGAGATT	295
QY	3090	AGTCTCTGGCTTTTGGGCGCTGATCACCGCATAGAAATGTTTACCAGAGGTGAGGAAATAA	3149
Db	296	AGCTCTGGCTTTTGGGCGCTGATCACCGCATAGAAATGTTTACCAG-----	340
QY	3150	TCACAAATTCATCATTGTGTGTAACCGTTGGACATATATTATGTGTGCTGCTTGT	3209
Db	341	-----	340
QY	3210	TGATTCGTATTTATATAGTCTTTTGATGGGCGAATGCTGAGATTTGGACTGCACAGCA	3269
Db	341	-----AGGTTTTGATGGGCGAAATCTGAAAGTTTGGACTGCACAGCA	364
QY	3270	GCTGAACCTGCGCAGCTTGCGTTGGTTTATGATTATACCTGTTGAACTATCTGTTGAC	3329
Db	385	GCTGAACCTGCGCAGCTTGCGTTGGTTTATGATTATACCTGTTGAACTATCTGTTGAC	444
QY	3330	AGTGTGACAGTCTCAGCAGATAGAAACCCGCTCTGTGTGGAACACTCGAGGAGTCT	3389
Db	445	AGTGTGACAGTCTCAGCAGATAGAAACCCGCTCTGTGTGGAACACTCGAGGAGTCT	504
QY	3390	GCTGTGCTATCTGATTTGGTTCATCCAGAAAACAATGCTACTGATTCAGAACTTACACA	3449
Db	505	GCTGTGCTATCTGATTTGGTTCATCCAGAAAACAATGCTACTGATTCAGAACTTACACA	564
QY	3450	ACAGATACGAAGTTTTTGTGCTCAAGTCAGGGTGAATAATCACTGAAGGCTCTGTTCTT	3509
Db	565	ACAGATACGAAGTTTTTGTGCTCAAGTCAGGGTGAATAATCACTGAAGGCTCTGTTCTT	624
QY	3510	GCATCATATATATCATATATGATAGCATGTCGAGCTTGGAGATTCTCTTTGTTCTGTA	3269
Db	625	GCATCATATATATCATATATGATAGCATGTCGAGCTTGGAGCTCTCTTTGTTCTGTA	684
QY	3570	ATTCTCTCTCTAAGTAGTATTAATAATGAACAAAAAATTAAGTTCTTGGCA	3626

Db	FEATURES	Source
Db	<p>RESULT 4 AI998415/c LOCUS DEFINITION thaliana Col-0, rosette-2 Arabidopsis ACCESSION AI998415 VERSION AI998415.1 KEYWORDS EST. SOURCE Arabidopsis thaliana (thale cress) ORGANISM Arabidopsis thaliana Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis. REFERENCE 1 (bases 1 to 561) Chen, J., Momiyama, M., Chan, E., Mooney, M., Carroon, B., Gilliland, D., Wang, X., Hillman, J., Guegler, K., Kim, C., Doyle, M., Brzoska, P., Gorgone, G., Burns, D., Griffin, J., Mouanoutou, M., Nguyen, D., Tan, R., Rose, M., Warren, B., Ton, B., Kastury, K., Borillo, C., Caprio, T., Policky, J., Suzuki, G., Argentine, C., Shah, S., Nobrega, A., Murry, L., Turner, C., Krikorian, S., Elder, L. and Hanson, D. Arabidopsis thaliana Gene Expression Microarray Unpublished (1999) Contact: David Smoller, Ph.D. Genome Systems, Inc., a wholly owned subsidiary of Incyte Pharmaceuticals, Inc. 4633 World Parkway Circle, St. Louis, MO 63134, USA Tel: 877-577-2733 Fax: 314-427-3324 Email: service@genomesystems.com. Location/Qualifiers 1..561 /organism="Arabidopsis thaliana" /mol_type="mRNA" /ecotype="Col-0" /db_xref="taxon:3702" /clone="701545606" /tissue_type="rosette" /dev_stage="4 - 7 weeks" /clone_lib="A. thaliana, Columbia Col-0, rosette-2" /note="Vector: pSPORT; Site_1: NotI; Site_2: SalI; cDNA library was derived from untreated rosette tissue from Arabidopsis thaliana, Columbia Col-0, at 4 - 7 weeks. Plants were grown in 1:1:1 peat moss/vermiculite/perlite soil at 22 deg. C +/- 3 deg. C under constant light, and watered with fertilizer. cDNA synthesis was initiated using a NotI-oligo(dT) primer. Double-stranded cDNA was blunt-ended, ligated to SalI adaptors, digested with NotI, size-selected, and cloned into the NotI and SalI sites of the pSPORT vector."</p>	<p>685 ATCTCTCTTAGATTAGTGTATTATATGACACAAAAATTACGTTCTTGCA 741</p>
Db	<p>Query Match 11.0%; Score 402.2; DB 1; Length 561; Best Local Similarity 82.2%; Pred. No. 2.4e-92; Matches 514; Conservative 0; Mismatches 20; Indels 91; Gaps 1;</p>	<p>AI998415 561 bp mRNA linear EST 08-SEP-1999 701545606 A. thaliana, Columbia Col-0, rosette-2 Arabidopsis thaliana cDNA clone 701545606, mRNA sequence.</p>
Db	<p>2970 TTTCTTCTTGGATTTTCTTCTTTCATTTAGGCTGACGAGCTGACATTCAGAG 3029 534 TTTCTTCTTGGATTTTCTTCTTTCATTTAGGCTGACGAGCTGACATTCAGAG 475 </p>	<p>3030 CACTTCCGAGATGGATGCTAGAGCTGACAGAGATATAGTATCCAGTGCAGAAATTA 3089 474 CACTTCCGAGATGGATGCTAGAGCTGACAGAGATATAGTATCCAGTGCAGAAATTA 415 3090 AGTCTTGGCTTTTGGGCTGATGACCGCATGAGAAATGTTACCAAGGTGAGGAATAA 3149 414 AGTCTTGGCTTTTGGGCTGATGACCGCATGAGAAATGTTACCAAGGTGAGGAATAA 370 3150 TCTACATTCATCAATGTTGTGTGAAAACTGTGTGACATGATTTATAGTGTGGCTGTTGT 3209</p>

FEATURES	SOURCE
1. 410	location/Qualifiers
/organism="Arabidopsis thaliana"	
/mol_type="mRNA"	
/db_xref="taxon:3702"	
/clone="RAFL07-83-O20"	
/dev_stage="rosette plants"	
/lab_host="DH10B"	

/clone_11b="RAF17"
/note="Site 1: BamHI; Site 2: SalI; subjected to
cold-treated (1, 2, 5, 10, 24 hr)"

ORIGIN

Query Match 10.3%; Score 376.4; DB 3; Length 410;
Best Local Similarity 99.5%; Pred. No. 1e-85;
Matches 388; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

3224 ATAGGTTTGGATGGCGGAATGCTGAAGATTGGACTGACAGACAGCGTGAACCTGCCA 3283
389 AGAGGTTTGGATGGCGGAATGCTGAAGATTGGACTGACAGACAGCGTGAACCTGCCA 330
3284 GCTTGGGTTGTTATGATTATATACACTGTGAACTATCTGTGACAGTGTGACAGTCTC 3343
329 GCTTGGGTTGTTATGATTATATACACTGTGAACTATCTGTGACAGTGTGACAGTCTC 270
3344 AGCAGATGGAACCCGCTGCTGTGTGGAAGCAACTCTGAGAGAGTCTGCTTCTATCTGA 3403
269 AGCAGATGGAACCCGCTGCTGTGTGGAAGCAACTCTGAGAGAGTCTGCTTCTATCTGA 210
3404 TTTGGTTTCATCCAGAAAACAATGCTACTGATGTGAGAACTTACACACAAAGATACGAAGT 3463
209 TTTGGTTTCATCCAGAAAACAATGCTACTGATGTGAGAACTTACACACAAAGATACGAAGT 150
3464 TTTCTGTCACAGTCAGGCTGTGGAAGCAACTCTGAGAGAGTCTGCTTCTATCTGA 3523
149 TTTCTGTCACAGTCAGGCTGTGGAAGCAACTCTGAGAGAGTCTGCTTCTATCTGA 90
3524 TCATATGATGATGTCTGAGCTTGGAGATTCTCTTTGTTCTGTAATTTCTCTCTAAG 3583
89 TCATATGATGATGTCTGAGCTTGGAGATTCTCTTTGTT- TGTAATTTCTCTCTAAG 31
3584 TTAGGTTTAAATGAACCAAAAAAATT 3613
30 TTAGGTTTAAATGAACCAAAAAAATT 1

RESULT 6

BP785511/c 415 bp mRNA linear EST 10-FEB-2005
LOCUS BP785511 RAF17 Arabidopsis thaliana cDNA clone RAF107-95-D11 3,
DEFINITION mRNA sequence.

ACCESSION BP785511
VERSION BP785511.1 GI:59261187

KEYWORDS

EST
Arabidopsis thaliana (thale cress)

SOURCE

Arabidopsis thaliana (thale cress)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

1 (bases 1 to 415)

Seki,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M., Akiyama,K.,
Iida,K., Enju,A., Sakurai,T., Arakawa,T., Carninci,P., Fukuda,S.,
Iida,J., Kawai,D., Sasaki,D., Shiraki,T., Hayashizaki,Y. and
Shinozaki,K.

REFERENCE

Large-Scale Analysis of RIKEN Arabidopsis Full-length cDNAs
Unpublished (2005)

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: meeki@rtc.riken.go.jp
An Arabidopsis full-length cDNA library was constructed essentially
as reported previously (Seki et al., 1998, 2002). This clone is in a
modified pBluescript vector.
Please visit our web site (http://pfweb.gsc.riken.jp and
http://large.gsc.riken.jp) for further details.
reversed clone.

FEATURES

Location/Qualifiers
1..415

/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAF107-95-D11"
/dev_stage="rosette plants"
/lab_host="DH10B"
/clone_11b="RAF17"
/note="Site 1: BamHI; Site 2: SalI; subjected to
cold-treated (1, 2, 5, 10, 24 hr)"

ORIGIN

Query Match 10.3%; Score 376.4; DB 3; Length 415;
Best Local Similarity 99.5%; Pred. No. 1e-85;
Matches 388; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

3224 ATAGGTTTGGATGGCGGAATGCTGAAGATTGGACTGACAGACAGCGTGAACCTGCCA 3283
389 AGAGGTTTGGATGGCGGAATGCTGAAGATTGGACTGACAGACAGCGTGAACCTGCCA 330
3284 GCTTGGGTTGTTATGATTATATACACTGTGAACTATCTGTGACAGTGTGACAGTCTC 3343
329 GCTTGGGTTGTTATGATTATATACACTGTGAACTATCTGTGACAGTGTGACAGTCTC 270
3344 AGCAGATGGAACCCGCTGCTGTGTGGAAGCAACTCTGAGAGAGTCTGCTTCTATCTGA 3403
269 AGCAGATGGAACCCGCTGCTGTGTGGAAGCAACTCTGAGAGAGTCTGCTTCTATCTGA 210
3404 TTTGGTTTCATCCAGAAAACAATGCTACTGATGTGAGAACTTACACACAAAGATACGAAGT 3463
209 TTTGGTTTCATCCAGAAAACAATGCTACTGATGTGAGAACTTACACACAAAGATACGAAGT 150
3464 TTTCTGTCACAGTCAGGCTGTGGAAGCAACTCTGAGAGAGTCTGCTTCTATCTGA 3523
149 TTTCTGTCACAGTCAGGCTGTGGAAGCAACTCTGAGAGAGTCTGCTTCTATCTGA 90
3524 TCATATGATGATGTCTGAGCTTGGAGATTCTCTTTGTTCTGTAATTTCTCTCTAAG 3583
89 TCATATGATGATGTCTGAGCTTGGAGATTCTCTTTGTT- TGTAATTTCTCTCTAAG 31
3584 TTAGGTTTAAATGAACCAAAAAAATT 3613
30 TTAGGTTTAAATGAACCAAAAAAATT 1

RESULT 7

B2437564/c 785 bp DNA linear GSS 13-DEC-2002
LOCUS B2437564 BONRN72TR BO.1.6_2_KB tot Brassica oleracea genomic clone BONRN72,
DEFINITION genomic survey sequence.

ACCESSION B2437564
VERSION B2437564.1 GI:26691135

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 785)
Ayale,M., Haas,B.J., Kumar,N., Wu,H., Xiao,Y., Van Aken,S.,
Utecherack,T.R., Wortman,J.R., White,O.R. and Town,C.D.
Whole genome shotgun sequencing of Brassica oleracea and its
application to gene discovery and annotation in Arabidopsis
Genome Res. 15 (4), 487-495 (2005)
15805490
Other_GSSs: BONRN72TF
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TR

FEATURES	Class:	sheared ends.
source	location/Qualifiers	
1..785		
/organism="Brassica oleracea"		
/mol_type="genomic DNA"		
/seq_id="OI000DH3"		
/db_xref="taxon:3712"		
/clone="BONRN72"		
/clone_id="EO_1.6-2_KB_tot"		
/note="Vector: pMOSt1; Site:1: total DNA inserted into pMOSt1 using BstXI linkers"		

[illegible]

RESULT 8	BP620404/c	434 bp	mRNA	linear	EST 26-JUN-2004
LOCUS	BP620404.1				
DEFINITION	BP620404.1 RAFL16 Arabidopsis thaliana cDNA clone RAFL16-37-Cl8 3', mRNA sequence.				
ACCESSION	BP620404.1				
VERSION	BP620404.1				
KEYWORDS	GI:49271586				
SOURCE	EST.				
	Arabidopsis thaliana (thale cress)				

REFERENCE	1 (bases 1 to 434)
AUTHORS	Seki, M., Narusaka, M., Kamiya, A., Ishida, J., Satou, M., Sakurai, T.,

TITLE •
Functional annotation of a full-length Arabidopsis cDNA collection
Science 296 (5565), 141-145 (2002)

COMMENT
Contact: Motoaki Seki
Motoaki Seki, Research Center

3-1-1 Koyradai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: meeki@rtc.riken.go.jp
reversed clone; Please visit our web site
(<http://pfweb.gsc.riken.go.jp/>) for further details

FEATURES	Location/Qualifiers
source	1. .434

```

/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAFL16-37-Cl8"
/lab_host="DH10B"
/clone_id="RAFL16"
/note="Site_1: BamHI, Site_2: SalI, dark-grown"

```

Query Match	9.8%	Score 359.4	DB 3	Length 434
Best Local Similarity	98.0%	Pred. No. 2.6e-81		
Matches 385	Conservative	0	Mismatches 6	Indels 2
				Gaps 2

3224 ATAGGTTTGGATGGCGAATGCTGAAGATTGGACTGACAGAGCAGCTGAACCTGCCA 3283

Db 402 AGAGGTTTGGATGGCGAATGCTGAAGATTGGACTGACAGAGCAGCTGAACCTGCCA 343

3284 GCCTGGGTTGCTTATGATTATACACTGTTGAAACTATCTGTTGACAGTGTGACAGTCTC 3343

Db 342 GCTTGGGTTGCTTATGATTATACACTGTGAACTATCTGTTACACTGTGACAGTCTC 283

3344 AGCAGATGGAACTCCGTCCTCTGGTGGAGCACTCTGGAGGAGTCTGCTGTCTATCTGA 3403

DB 282 AGCAGATGGAACTCGCTCTGGTGGAGCAACTCTGGAGGAGCTGCTTGCTAATCTGA 223

3404 110311C41C4A004C4A1B1AC1B41B1C4A04C1AC4A1C4A041AC4A041 3405

2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020 2021 2022 2023 2024 2025 2026 2027 2028 2029 2030 2031 2032 2033 2034 2035 2036 2037 2038 2039 2040 2041 2042 2043 2044 2045 2046 2047 2048 2049 2050 2051 2052 2053 2054 2055 2056 2057 2058 2059 2060 2061 2062 2063 2064 2065 2066 2067 2068 2069 2070 2071 2072 2073 2074 2075 2076 2077 2078 2079 2080 2081 2082 2083 2084 2085 2086 2087 2088 2089 2090 2091 2092 2093 2094 2095 2096 2097 2098 2099 2100 2101 2102 2103 2104 2105 2106 2107 2108 2109 2110 2111 2112 2113 2114 2115 2116 2117 2118 2119 2120 2121 2122 2123 2124 2125 2126 2127 2128 2129 2130 2131 2132 2133 2134 2135 2136 2137 2138 2139 2140 2141 2142 2143 2144 2145 2146 2147 2148 2149 2150 2151 2152 2153 2154 2155 2156 2157 2158 2159 2160 2161 2162 2163 2164 2165 2166 2167 2168 2169 2170 2171 2172 2173 2174 2175 2176 2177 2178 2179 2180 2181 2182 2183 2184 2185 2186 2187 2188 2189 2190 2191 2192 2193 2194 2195 2196 2197 2198 2199 2200 2201 2202 2203 2204 2205 2206 2207 2208 2209 2210 2211 2212 2213 2214 2215 2216 2217 2218 2219 2220 2221 2222 2223 2224 2225 2226 2227 2228 2229 2230 2231 2232 2233 2234 2235 2236 2237 2238 2239 2240 2241 2242 2243 2244 2245 2246 2247 2248 2249 2250 2251 2252 2253 2254 2255 2256 2257 2258 2259 2260 2261 2262 2263 2264 2265 2266 2267 2268 2269 2270 2271 2272 2273 2274 2275 2276 2277 2278 2279 2280 2281 2282 2283 2284 2285 2286 2287 2288 2289 2290 2291 2292 2293 2294 2295 2296 2297 2298 2299 2300 2301 2302 2303 2304 2305 2306 2307 2308 2309 2310 2311 2312 2313 2314 2315 2316 2317 2318 2319 2320 2321 2322 2323 2324 2325 2326 2327 2328 2329 2330 2331 2332 2333 2334 2335 2336 2337 2338 2339 2340 2341 2342 2343 2344 2345 2346 2347 2348 2349 2350 2351 2352 2353 2354 2355 2356 2357 2358 2359 2360 2361 2362 2363 2364 2365 2366 2367 2368 2369 2370 2371 2372 2373 2374 2375 2376 2377 2378 2379 2380 2381 2382 2383 2384 2385 2386 2387 2388 2389 2390 2391 2392 2393 2394 2395 2396 2397 2398 2399 2400 2401 2402 2403 2404 2405 2406 2407 2408 2409 2410 2411 2412 2413 2414 2415 2416 2417 2418 2419 2420 2421 2422 2423 2424 2425 2426 2427 2428 2429 2430 2431 2432 2433 2434 2435 2436 2437 2438 2439 2440 2441 2442 2443 2444 2445 2446 2447 2448 2449 2450 2451 2452 2453 2454 2455 2456 2457 2458 2459 2460 2461 2462 2463 2464 2465 2466 2467 2468 2469 2470 2471 2472 2473 2474 2475 2476 2477 2478 2479 2480 2481 2482 2483 2484 2485 2486 2487 2488 2489 2490 2491 2492 2493 2494 2495 2496 2497 2498 2499 2500 2501 2502 2503 2504 2505 2506 2507 2508 2509 2510 2511 2512 2513 2514 2515 2516 2517 2518 2519 2520 2521 2522 2523 2524 2525 2526 2527 2528 2529 2530 2531 2532 2533 2534 2535 2536 2537 2538 2539 2540 2541 2542 2543 2544 2545 2546 2547 2548 2549 2550 2551 2552 2553 2554 2555 2556 2557 2558 2559 2560 2561 2562 2563 2564 2565 2566 2567 2568 2569 2570 2571 2572 2573 2574 2575 2576 2577 2578 2579 2580 2581 2582 2583 2584 2585 2586 2587 2588 2589 2590 2591 2592 2593 2594 2595 2596 2597 2598 2599 2600 2601 2602 2603 2604 2605 2606 2607 2608 2609 2610 2611 2612 2613 2614 2615 2616 2617 2618 2619 2620 2621 2622 2623 2624 2625 2626 2627 2628 2629 2630 2631 2632 2633 2634 2635 2636 2637 2638 2639 2640 2641 2642 2643 2644 2645 2646 2647 2648 2649 2650 2651 2652 2653 2654 2655 2656 2657 2658 2659 2660 2661 2662 2663 2664 2665 2666 2667 2668 2669 2670 2671 2672 2673 2674 2675 2676 2677 2678 2679 2680 2681 2682 2683 2684 2685 2686 2687 2688 2689 2690 2691 2692 2693 2694 2695 2696 2697 2698 2699 2700 2701 2702 2703 2704 2705 2706 2707 2708 2709 2710 2711 2712 2713 2714 2715 2716 2717 2718 2719 2720 2721 2722 2723 2724 2725 2726 2727 2728 2729 2730 2731 2732 2733 2734 2735 2736 2737 2738 2739 2740 2741 2742 2743 2744 2745 2746 2747 2748 2749 2750 2751 2752 2753 2754 2755 2756 2757 2758 2759 2760 2761 2762 2763 2764 2765 2766 2767 2768 2769 2770 2771 2772 2773 2774 2775 2776 2777 2778 2779 2780 2781 2782 2783 2784 2785 2786 2787 2788 2789 2790 2791 2792 2793 2794 2795 2796 2797 2798 2799 2800 2801 2802 2803 2804 2805 2806 2807 2808 2809 2810 2811 2812 2813 2814 2815 2816 2817 2818

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

[illegible]

[illegible][illegible]

RESULT 9
CX195205 624 bp mRNA linear EST 28-DEC-2004
LOCUS CX195205
DEFINITION 13-B019365-021-009-J03-T7R ADIS-MP1Z 021 Brassica napus cDNA clone
MP1ZP102220390, mRNA sequence.
ACCESSION CX195205
VERSION CX195205.1 GI:56842629
KEYWORDS EST.
SOURCE Brassica napus (rape)
ORGANISM Brassica napus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1 (bases 1 to 624)
AUTHORS Jakoby, M., Lehmann, D. and Weisshaar, B.
TITLE direct submission to GenBank (ADIS-MP1Z 021)
JOURNAL Unpublished (2004)
COMMENT Contact: Bernd Weisshaar
Chair of Genome Research
Bielefeld University, Institute for Genome Research
Universitaetsstrasse 25, D-33594 Bielefeld, Germany
Tel: +49-521-106 6873
Fax: +49-521-106 6423
Email: bernd.weisshaar@uni-bielefeld.de
Insert length: 624 Std Error: 0.00
Plate: 9 row: J column: 3
Seq primer: T7R CTATACGACTACTATAGGA.
Location/Qualifiers
1. 624
/organism="Brassica napus"
/mol_type="mRNA"
/cultivar="Expres 617"
/db_xref="GABI:1111775"
/db_xref="taxon:3708"
/clone="MP1ZP10220390"
/tissue_type="flowers and flower buds"
/dev_stage="young flowers and flower buds"
/clone_1db="ADIS-MP1Z 021"
/note="Vector: pSPOR1; Site_1: SalI; Site_2: NotI; cDNA
library from Brassica napus, strain Expres 617; RNA was
from young flowers and flower buds of two greenhouse
plants; library was made at the Max-Planck-Institute for
Plant Breeding Research, Cologne, Germany; cloning sites
Sali-NotI, primer sites and orientation:
T7-Sali-CCAGCGGTCG-5'utr-cDNA-polyA-CC-NotI-SP6; made by
Marc Jakoby 09/2000; PI: Bernd Weisshaar. Sequence
submission managed by R2PD/GABI-Primary database:
http://gabi.rzpd.de"

ORIGIN
Query Match 9.8%; Score 359.2; DB 8; Length 624;
Best local Similarity 76.9%; Pred. No. 3.3e-81;
Matches 518; Conservative 0; Mismatches 93; Indels 63; Gaps 4;
Query 2035 AGGAATCAGCTATGTTGGAGTTGTTTGAAGATTCAATCGTATGATCATATGAT 2094
2 AGGAATCAGCTATGTTGACTTTGCTTTGAGAAATTAATCGATGACATACGAC 61
Db 2095 CTCCCTGACTATGCAAAATTTGGAACTGTGTTGGAGGAGGTTGCTTTCTTAGGTTG 2154
62 CTCCCTGAGCTCTGCAAAATTAATTTGAGACCTGTGTTGGAGGAGGTTGCTTTCTTAGATTG 121
Query 2155 AGAGACACCAAGATTAATAATTTAACTCGGGAGCTACTATGATGATCGATGCTTTG 2214
122 AGAGATACCAAGATTAATAATTTAACTCGGGAGTTATTAAGATGCCATGGTTTG 181
Query 2215 AGTTACTTGAAGAGTGAAGTGAAGTGAAGGTTCTCTTTAGCTGTGCTGCAACTATG 2274
182 AGTTACTTGAAGAGTGAAGTGAAGTGAAGGTTCTCTTTAGCTGTGCTGCAACTATG 241
Query 2275 GCAGAGATTGAGCCGAGCATGTGAAGCTATGCTATGCAAGGCACTGCAGAAAGTTT 2334

Db 242 GCAGAGATTGAGCCGAGCATGTGAAGCTATGCTATGCAAGGCACTGCAGAAAGTTT 301
Query 2335 CCTTCCGCTATACAGATAGAACTCGGCTGAACCCAGATGTGCAAGACAGCTGTT 2394
Db 302 CCTTCTGCTATTCA-----GATCGCAAGACAGCGTGCT 337
Query 2395 AGTGTAGATCTGTGTGTAACATGATAGCCGTGA---TGTGAGCCGTGTGCTTAAT 2451
Db 338 ACTGTAGATCATGTGTGTAACATGATGATGATGATGATGATGATGATGATGATGAT 397
Query 2452 GCAGAGCTGTGAACCCCTGTGAACACTTTGAACCTAATGATGATGATGATGATGATG 2511
Db 398 GCAGCAACTGGAAGAGATCCGAATCTTTGATA----- 431
Query 2512 GTCTCAGAGATGACGTTGATGAACACTGTTGAATGTCCTGTTGCTGATGATGATGA 2571
Db 432 ----CAAGTCTACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 487
Query 2572 GAGGCAAGTGTGAAGATCTAGCTGTGTGTGCAATTGCACTGATTTCACTGTCAGC 2631
Db 488 GAGGCAAGTGTGAAGATCTCTGCGCGGTGTGTGTGGAATAATTTGCTGCGCAGC 547
Query 2632 CAGAAGTATTTCTTAA-----GAGCTCATCTTTTCAAGCCAGAGATAGGTTCT 2685
Db 548 CTGAGGATATGCTATTAAGGCGGAGCTCATGTTGCAACGCAAGATATGATCTCC 607
Query 2686 TCTATGAACTGA 2699
Db 608 TCTACGCACTGA 621

RESULT 10
CD573714 897 bp mRNA linear EST 12-JUN-2003
LOCUS UCRPT01_01_P12.T3 Poncirus trifoliata CTV-challenged cDNA library -
DEFINITION UCR Poncirus trifoliata cDNA clone UCRPT01_01_P12, mRNA sequence.
ACCESSION CD573714
VERSION CD573714.1 GI:31669616
KEYWORDS EST.
SOURCE Poncirus trifoliata
ORGANISM Poncirus trifoliata
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Sapindales; Rutaceae; Poncirus.
REFERENCE 1 (bases 1 to 897)
AUTHORS Roose, M.L., Ye, X., Federici, C.F., Close, T.J., Fenton, R.D.,
Wanamaker, S., Choi, Y. and Kingan, T.
TITLE Development of EST Resources and New Genetic Markers for California
Citrus - Poncirus trifoliata CTV-challenged phloem - UCR
JOURNAL Unpublished (2003)
COMMENT Contact: Mikeal Roose
Department of Botany & Plant Sciences, University of California
Riverside, CA, 92521-0124, USA
Tel: 909/8744137
Fax: 909/8744137
Email: mikeal.roose@ucr.edu
Seq primer: T3.
Location/Qualifiers
1. 897
/organism="Poncirus trifoliata"
/mol_type="mRNA"
/cultivar="Pomeroy OP"
/db_xref="taxon:37690"
/clone="UCRPT01_01_P12"
/tissue_type="phloem"
/dev_stage="10 - 30 cm shoots"
/lab_host="E. coli TUC121"
/clone_1db="Poncirus trifoliata CTV-challenged cDNA
library - UCR"
/note="Vector: Lambda Uni-ZAP XR, excised phagemid;
Site_1: EcoRI; Site_2: XhoI; Plants were grown in the
greenhouse at University of California, Riverside, the

scion was a open-pollinated (very probably selfed) seedling of *Poncirus trifoliata* cv Pomeroy that was selected as homozygous for the Ctv resistance gene. The rootstock was sweet orange infected with citrus tristeza virus (CTV) isolate T514 over 1 year before sampling (CTV infects sweet orange, but not genotypes carrying the Ctv resistance gene. Shoots 10-30 cm long were harvested in October 2000, and the green phloem (bark) was removed and frozen quickly in dry ice. Total RNA was extracted using Trizol reagent (Gibco). Poly(A) RNA was purified, a cDNA library was made, and 0.5 million primary lambda cDNA clones were in vivo excised to give a population of phagescript SK(-) phagemids. All steps to this point were performed in the M. Roese lab at the University of California, Riverside by X. Ye. Phagemids were plated, plasmid DNA purified, cDNA clones archived, and DNA sequences determined bi-directionally using an ABI3730 at the University of California Riverside Genome Institute, Core Instrumentation Facility, (Choi, Kingan). Chromatogram files were transmitted to UC Riverside (by Choi), then processed at UC Riverside (by Manamaker) using the Harvester pipeline (<http://harvest.ucr.edu>) to remove vector and cloning oligo sequences and various contaminants, and to trim to a high quality region. Sequences that retained a phred 17 region of at least 100 bases were deposited to GenBank."

ORIGIN

Query Match 9.2%; Score 336.4; DB 6; Length 897;
Best Local Similarity 63.9%; Pred. No. 2.9e-75;
Matches 574; Conservative 0; Mismatches 241; Indels 83; Gaps 1;
QY 1391 GCAAGTACCTTGCACCGGATTTACGTCAGAAATTGATGAGCTTTGGAGAGATCACT 1450
Db 80 GCCAGTGTCTCGCTCCAGATTTTGACGACAGANTGACGAGCTTGAGAGATCTAC 139
QY 1451 CCGGCTTATGCTTGAGACTACTTGCGCTTACCGCTTGATGATACGCTGCGAAAAGA 1510
Db 140 CCAGCGCTGTCTGGAACCTTTAGGCTTACCCCTTACTGTGATGATCAAGCAAGAAGA 199
QY 1511 CTAATGCTTAAAGCGGTGCGGAATTTTGTGCTGTGTGAGAGAGGTGAGCATCA 1570
Db 200 GAAAGAGACTTCATGATGATCTCAACACTCACTGATGAGAGGAGGAGCCACT 259
QY 1571 GCTCTTGTGGGGGTTTGACCCGTCAGAGTTTGAATGAGCGCTTTTACGATGCA 1630
Db 260 GCAATGTGCGGGGATTTACAGGTAATCTTTCATGATGAGGCTTTTCAAGTATGCA 319
QY 1631 GCTGCTGACAGGTATACAGTTTAAATTTCTTAAATTTCTTAAAGATATTAAT 1690
Db 320 TCAGCTGAGC----- 329
QY 1691 TTAGGTTTCTATTTATGATGTTGTGTGAGGTGATCTTTTGTAGCTACCCCA 1750
Db 330 -----AGTTTAACTATTTTCTCCACACCA 356
QY 1751 GCAATATTCAGCAGAGTCATTTGAAGTTTGAAGTTGCACTTGTGCTGCTCAAG 1810
Db 357 ATAGATCCAGCTGAACCTTTGAAGCTTATGAGTGGCATTGCACTTTTGTCCCAAG 416
QY 1811 CTTTATTTGTTAAGACCACTTTTACAGATGCTGATPAAGCAATTCAGCAACTTC 1870
Db 417 CTTTGTGGGTAAAGCACTCATCTCATGTCAGATGCTGATPAACATGTTCAAGCACTTC 476
QY 1871 AGCAGGCTAAGTATGCTATGAGATTTCTGCGATTTGTATGATACAGGAATATT 1930
Db 477 AGCAGAAATTAAGTACACAGCTTTACGGGACCTTGTCTATATTAATTCCTTGAAGAAAC 536
QY 1931 GGGAGATAGACTTGCTGCTAGAAAAGGAGCTCTGTCAGCTGCTTATAGGCAAGTTGATG 1990
Db 537 ATGAGATGAGATTTGCAATGGAAGGGGTCTATGTCTCTGCTTGTAGTATAGCTTGATG 596
QY 1991 AATGCCGTATGTGTTGGGCTTTAAGCAGTGAAGATTCACAAATATGAAATCCAGCTATTG 2050

|||||
Db 597 AGTCCCTTATGATGTTGGGCTTATGATGATTAAGTCCCATACAGAAATCCAGCCATTG 656
QY 2051 TGGAGTTGTTTGGAGAAATTCAAATGCGATGACAAATGATATTCCTCCGAGCTATGCA 2110
Db 657 TAGTTTGTGTTGGAAACTCAAGAGAACTATGACAAATGATCTTCTGTTCTGTGA 716
QY 2111 AATGTTGAACCTGTTGGAGAGGGGTTGCTTCTTCTAGTTTCAGAGACCAAGATGA 2170
Db 717 AGCTGTAGAGACATGCTGCGAGAGGTTGTTTCCCTAGATTCAAGGACACCAAGGATA 776
QY 2171 AAAATTAACTCGGGGACTTACTATGATATCTTATGCTTTGAGTTTACTTGGAAAG 2230
Db 777 TACGTTCAAACTGGGTGATTAATGATATCTTATGCTTCCCTTCGATTTTGAAGAAC 836
QY 2231 TGGAGTAGTTCAAGGGTTCTCTTATGCTGCTGCAATATGCAAGAGATTGAGAC 2288
Db 837 TGGAGGAATGATGATGTTCCCTTACTGACAGCAGCTGCAATATGTTAGGCTTGAGC 894

RESULT 11
DU050213/c 871 bp DNA linear GSS 12-AUG-2005
LOCUS 144781 tomato HindIII BAC Library Lycopersicon esculentum genomic
DEFINITION clone LE_HBa0013B21 5, genomic survey sequence.
ACCESSION DU050213
VERSION DU050213.1 GI:72480609
KEYWORDS GSS.
SOURCE Lycopersicon esculentum (Solanum lycopersicum)
ORGANISM Lycopersicon esculentum
Bakaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamiales; Solanales; Solanaceae; Solanum; Lycopersicon.

REFERENCE
AUTHORS Mueller L.A., Buels R.M., Wang, Y., Tanksley, S.D., Giovannoni, J.J.,
Van Eck, J. and Stack, S.
BAC end sequencing from three Solanum lycopersicon libraries
Unpublished (2005)
Other GSSs: 144780
COMMENT
Contact: Lukas Mueller
Tanksley Lab, Dept. of Plant Breeding
Cornell University
251 Emerson Hall, Ithaca, NY 14853, USA
Tel: 607-255-6557
Fax: 607-255-6683
Email: sgn-feedback@cornell.edu
Insert length: 76680 Std Error: 0.00
Plate: 13 Row: B Column: 21
Seq primer: T7
Class: BAC ends
High quality sequence start: 43
High quality sequence stop: 789.
Location/Qualifiers

FEATURES
source
1..871
/organism="Lycopersicon esculentum"
/mol_type="genomic DNA"
/cultivar="Heinz 1706"
/db_xref="taxon:4081"
/clone="LE_HBa0013B21"
/lab_host="E. coli"
/clone_id="Tomato HindIII BAC library"
/note="Vector: pBelobAC11, site_1: HindIII"

ORIGIN

Query Match 9.1%; Score 334; DB 10; Length 871;
Best Local Similarity 67.4%; Pred. No. 1.2e-74;
Matches 503; Conservative 0; Mismatches 235; Indels 8; Gaps 2;
QY 1374 TTGTTGTTAGAGAGAGAGCAAGTAGCTTGCACCGGATTTTACGTGCAAAATGATGAGA 1433
Db 791 TTTTCCAGGAAGAGAGGCAAGATATCTTGCCTTGTATCTGCAATCCCAATGATGATAGA 732
QY 1434 CTTTGAAGAGATCACTCGCGTTATGTTGAGAGCTTATGCGCTTACCGCTTGATG 1493

Db	731	CATTGGAAGATTAATCAGGCTACGTACTTGAACTTCTAGCTTTCTCTTGATGATG	672
Oy	1494	ATTACGCTGCGAAAAGACTAAATGGTTTAAAGCGATGTCGCGAATATTTTGTGCTGTGTG	1553
Db	671	AATACCGAATGAAAAGAGTAGAGCTCTTCAAGGTGTGCGCAACATTTTGTGGCTGTGTG	612
Oy	1554	GAGAGGGTGGAGCATCAGCTCTGTGTGGGGGGTTTAAACCCGTGAGAAATTATGATGAG	1613
Db	611	GAGGAGGAGGAGCGACTCGCAATTTACGAGGGGGTTTACACGAGMAATTTCAATGAATVAGG	552
Oy	1614	CGTTTTTACGATGACAGCTGTGTGACAGS-----TATACAGTTTATGATTAACCTTTT	1668
Db	551	CCTTCCTACGAGATGACAGCTGTGTGACAGGTTGGCTTAACTTCAATTAATTTTGGTT	492
Oy	1669	ATTCTTTAGCATGATATPACTTTAGTTTCTCATTTTAAATGATGTTGTGTAGGTT	1728
Db	491	TCTTTTGTAACTGTTCTTTTCACTCTCTCTCTCATTTTCCCTCTCTTTTAAAAAGTGG	432
Oy	1729	GATCTTTTGTAGCTAACCCCAAGCAATTTCCAGAGAGTCAATTGAAAGTTTACGAATTT	1788
Db	431	GACCTCTTGTGCGAAGCGCAAGTAAACATCTCTGAGAAAGCTTTGAAAGTTTACGGGGTGG	372
Oy	1789	GCATTTGCTCTGTGGCTCAAGCTTTTATTTGGTAAAGACCAACCTTTTACAGAGTCT	1848
Db	371	GCGGTGACCTTGTGTGCTCAAGCTTTTGTGGAAAAAAACCTCACTCATTCAGAGATCT	312
Oy	1849	GATTAAGCAATTCACGAACCTTACAGGCTTAAAGTAAATGGCTAATGGAAGATTCCTGCGATG	1908
Db	311	GATTAACCTTTTTCAGACAGCTTACAGACCAAAAGTTTACAGCTTACGGCAGTTCTGTGCT	252
Oy	1909	TTGATATGATACACGGAATATTTGGAGATTAACCTTCGCTCTTAAAGAAAGGGGACTCTGTGCA	1968
Db	251	GTATACACTGTTTAAAGAAAACCGTAAATTAAGCTTCGCTTTGGAAAGGGGCTTTGTGTCA	192
Oy	1969	CTGCTTATAGGCAAGTATGATGAATCCGTAATGTGCTTGGGCTTTAGACATGAGGATTTCA	2028
Db	191	CTGCTGTGTGAGAAAGTGAATGATGATGTGCTTGTGTTGGGCTCTAGAACATGAGGACCTCC	132
Oy	2029	CAATATAGGAATCCAGCTAATTTGGAGTTTGTTTTGGAAATTCAAATCGTATGATC---	2085
Db	131	CTTATATAGATTCAGCTATATGTGACTTTTGTGTGAGAAACACTMAAGAGACGAACACGA	72
Oy	2086	AATGATGATCTCCCTGAGACTAATGCA 2111	
Db	71	AATGATCTGCTCCCTGAGACTAATGTA 46	

RESULT 12	
B0834167	
LOCUS	B0834167
DEFINITION	B0834167 534 bp mRNA linear EST 14-MAR-2003
ACCESSION	AL1870128 Arabidopsis lyrata Inference PCW-PER Library
VERSION	B0834167
KEYWORDS	B0834167.1 GI:28951482
SOURCE	EST.
ORGANISM	Arabidopsis lyrata
	Arabidopsis lyrata
	Eunaryota, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
	rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE	1 (bases 1 to 534)
AUTHORS	Barrier,M., Bustamante,C.D., Yu,J. and Purnganan,M.D.
TITLE	Selection on rapidly evolving proteins in the Arabidopsis genome
JOURNAL	Genetics 163 (2), 723-733 (2003)
PUBMED	12618409
COMMENT	Contact: Barrier M

North Carolina State University
3513 Gardner Hall, Box 7614, Raleigh, NC 27695, USA
Tel: 919 515 1761
Fax: 919 515 1695
Email: mbarric@unity.ncsu.edu

```

FEATURES
    source
        Plate: 1 row: H column: 5
        Seq primer: T3
        Location/Qualifiers
            1..534
                /organism="Arabidopsis lyrata"
                /mol_type="mRNA"
                /culturvar="Karahumaki"
                /db_xref="taxon:59689"
                /clone_id="PIWB1-D03"
                /tissue_type="Inflorescence"
                /clone_lib="Arabidopsis lyrata Inflorescence pCMV-PCR
                Library"
                /note="Vector: pCMV-PCR (Stratagene) : Created using PCR
                Library Construction kit (Stratagene) "

```

Query Match		8.6%;	Score 315.8;	DB 5;	Length 534;
Best Local Similarity		76.9%;	Pred. No. 5,4e-70;		
Matches	442;	Conservative	0;	Mismatches	42;
				Indels	91;
				Gaps	1;
Qy	2971	TTCTTTCCTTGAAATTTTTCCTTTTCGATTTAGGCTGACGACGCTGACGATTCAGAGC	3030		
Db	51	TTCTTCATGGAATCTGACATTTCTACATAGGGTCTGAGACCTGATTCAGAGC	110		
Qy	3031	ACTTCCAGAAATGAGTCTGAGCTGAGAAATATATGTATCCAGTGGAGAAAGATTAA	3090		
Db	111	ACTTCCAGATGAGTCTGAGCTGAGAGAGTATATGATCCAAATGGAGAAAGATCAA	170		
Qy	3091	GTCCTGGCTTTTGGGCTGATACCGCATAGAAATTTTACCAGAGGTGAGGGAATTAAT	3150		
Db	171	GTCACAGCTTTTGGGATATCACTGCAATAGAAATCTACAG-----	214		
Qy	3151	CTACAAATTCATCAATTTGTGTGAAACTGTGACATGATTAATATGTCTGTCCCTGTTT	3210		
Db	215	-----	214		
Qy	3211	GATTCGTATTATATAGGTTTGGATGGGGAATGCTGAGATTTTGGACTGACAGACAG	3270		
Db	215	-----AGTTTTGATGGGGGAATGCTGAAAGATTTGGACTGACAGACCTG	259		
Qy	3271	CTGAAACTGGCAGCTTTGGGTGTTATGATTATACACTGTGTGAAACTATCTGTTGACA	3330		
Db	260	TGGAACCTGGCGAGCTTTGGCTGGTTATATATATACACTGTTGAAACTATCTGTTGACA	319		
Qy	3331	GTCGTACAGTCTCAGCAGATGGAACCCGTGCTCTGTGTGAAGCACTCTGAGAGATCTG	3390		
Db	320	GTCGTACAGTCTCAGCAGATGGAACCCGTGCTCTGTGTGAAGCAACTTGTGAGAGATCTG	379		
Qy	3391	CTTGTCTATCTGATTTGGTTCATCCAGAAAACAATGCTACTGATGTGAGAACTACACAA	3450		
Db	380	CTTGTCTATCTGATTTGGTTCATCCAGAAAACAATGCTACTGATGTGAGAACTACACAA	439		
Qy	3451	CAAGATACGAAGTTTTCTGTGCCAAGTCAGAGGTGAGAAATCACTGAAGGCTCTGTTCCTG	3510		
Db	440	CAAGATACGAAGTCTGTGTGCCAAGTCAGAGGTGAGAAATCACTGAAGGCTCTGTTCCTG	489		
Qy	3511	CATCATATATTAATCATATGTACAGATGTCTGAACT	3545		
Db	500	CATCATATATATGCTTTATATATACATACCCGAACT	534		

RESULT	13
LOCUS	AV830764
DEFINITION	320 bp mRNA linear EST 01-APR-2002
ACCESSION	AV830764 RAFLY Arabidopsis thaliana cDNA clone RAFL09-76-G11 5', mRNA sequence.
VERSION	AV830764
KEYWORDS	AV830764.1 GI:19672824 EST.
SOURCE	Arabidopsis thaliana (thale cress)
ORGANISM	Arabidopsis thaliana Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;

REFERENCE	
AUTHORS	rosides; eurosids II; Brassicales; Brassicaceae; Arabidopsis. 1 (bases 1 to 320) Seki,M., Narusaka M., Ishida,J., Kamaya,A., Saitou,M., Nakajima,M., Oono,Y., Sakurai,T., Carninci,P., Kawai,J., Itoh,M., Ishii,T., Arakawa,T., Shibata,K., Shinagawa,A., Muramatsu,M., Hayashizaki,Y. and Shinozaki,K.
TITLE	Large scale analysis of Arabidopsis full-length cDNA (2002b)
JOURNAL	Unpublished (2002)
COMMENT	Contact: Motoaki Seki Plant Functional Genomics Research Group RIKEN Genomic Sciences Center 3-1-1 Koyadaai, Tsukuba, Ibaraki 305-0074, Japan Tel.: 81-298-36-4359 Fax: 81-298-36-9060 Email: msek@rc.riken.go.jp An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). cDNA cleaved with BamHI and XhoI was ligated to modified lambda FL-C-1 vector (Carninci et al., submitted for publication) digested with BamI and SalI. This clone is in a modified pBluescript vector. Please visit our web site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further details.
FEATURES	
source	Location/Qualifiers 1..320 /organism="Arabidopsis thaliana" /mol_type="mRNA" /db_xref="taxon:3702" /clone="RAFL09-76-G11" /dev_stage="plants at various developmental stages from germination to mature seeds" /lab_host="DH10B" /clone_id="RAFL9" /note="Site 1: BamHI, Site 2: SalI, subjected to dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24 hr) treatments"
ORIGIN	
Query Match	8.5%; Score 311.8; DB 1; Length 320;
Best Local Similarity	98.4%; Pred. No. Se-69;
Matches 313; Conservative	0; Mismatches 5; Indels 0; Gaps 0;
Db	
QY	368 GATTTAACCTTACTACTCAAAATCAAATTCATTAACCCGTAGACGCAAAAGCTGC 427
	2 GATTTAACCTTACTACTCAAAATCAAATTCATTAACCCGTAGACGCAAAAGCTGC 61
QY	428 TTCAATAATGTAAACAGAACAAGTTTGTGGTAGTGCTTAAGAACACTGCCATGAAG 487
Db	62 TTCAATAATGTAAACAGAACAAGTTTGTGGTAGTGCTTAAGAACACTGCCATGAAG 121
QY	488 CTCGAGTCACGTCGGCATTTGGTCTCTCCCAATTCATTTAAGCGATTACACCGGCGA 547
Db	122 CTCGAGTCACGTCGGCATTTGGTCTCTCTCCCAATTCATTTAAGCGATTACACCGGCGA 181
QY	548 CGACAAAGCTCCGAGTACGCAACAACCTCTCAACTATCTGCTCCGCCAGCAAAATGGG 607
Db	182 CGACAAAGCTCCGAGTACGCAACAACCTCTCAACTATCTGCTCCGCCAGCAAAATGGG 241
QY	608 CCGACCGCTTCTCTCGCATTTCAATTTCACTCGGATTCTCTCTCTCTCTTGGCGCA 667
Db	242 CCGACCGCTTCTCTCGCATTTCAATTTCACTCGGATTCTCTCTCTCTCTTGGCGCA 301
QY	668 CCGCACACGACGACGCGCA 685
Db	302 CCGCACACGACGACGCGCA 319
RESULT_14	
BZ765587	320 bp DNA linear GSS 13-MAR-2003
LOCUS	SALK_131837.28.70.x Arabidopsis thaliana TMDNA insertion lines
DEFINITION	Arabidopsis thaliana genomic clone SALK_131837.28.70.x, genomic survey sequence.
ACCESSION	BZ765587

VERSION	B2765587.1	GI:28938140
KEYWORDS	GSS.	
SOURCE	Arabidopsis thaliana (thale cress)	
ORGANISM	Arabidopsis thaliana	
REFERENCE	Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosoid; eurosoid II; Brassicales; Brassicaceae; Arabidopsis.	
AUTHORS	1 (bases 1 to 320) Alonso, J.M., Leisbe, T.J., Barajas, P., Chen, H., Cheuk, R., Gadrinab, C., Jeike, A., Karnes, M., Kim, C.J., Parker, H., Prednis, L., Shum, P., Zimmerman, J. and Ecker, J.R.	
TITLE	A sequence-indexed library of insertion mutations in the Arabidopsis Genome	
JOURNAL	Unpublished (2001)	
COMMENT	Contact: Joseph R. Ecker Salk Institute Genomic Analysis Laboratory (SIGNAL) The Salk Institute for Biological Studies 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA Tel: 858 453 4100 x1752 Fax: 858 558 6379 Email: j.ecker@salk.edu This is single pass sequence recovered from the left border of TDNA. This sequence lies within 300 bases of the 5' end of AC5942480.	
FEATURES	Class: TDNA tagged. Location/Qualifiers 1..320 /organism="Arabidopsis thaliana" /mol_type="genomic DNA" /ecotype="Col-0" /db_xref="taxon:3702" /clone_1fb="Arabidopsis thaliana TDNA insertion lines" /note="TPCR was performed on Arabidopsis thaliana lines each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://signal.salk.edu/tcna_protocol.html "	
ORIGIN		
Query Match	8.4%; Score 307.2; DB 9; Length 320;	
Best Local Similarity	97.5%; Pred. No. 7.8e-68;	
Matches 312; Conservative 0; Mismatches 8; Indels 0; Gaps 0;		
QY	343 TTTTCAGTACGTGATGATTTTGTGATTAACCTAATCAAAATTCAT	402
Db	1 TTTTCAGTACGTGATGATTTTGTGATTAACCTAATCAAAATTCAT	60
QY	403 AAACCCCTAGACGACCAAGAGCTCTCAATATGTAACAGAACTTTTGTAGTA	462
Db	61 AAACCCCTAGACGACCAAGAGCTCTCAATATGTAACAGAACTTTTGTAGTA	120
QY	463 GCCTAAAGACACTCCCATGGAAGCTCTGAGTCACGTGGCAATGGCTCTCCCATTC	522
Db	121 GCCTAAAGACACTCCCATGGAAGCTCTGAGTCACGTGGCTCTCCCATTC	180
QY	523 CAATTATGCCGATTAACACCGGCGACGACAAAGCTCCGAGTACGACCAACCTCTACA	582
Db	181 CGATTATGCCGATTAACACCGGCGACGACAAAGCTCCGAGTACGACCAACCTCTACA	240
QY	583 ACTATCGCTCCGCGACGAATGGGCGGACCGTCTTCTCCGACTCAATTTCACTCC	642
Db	241 ACTATCGCTCCGCGACGAATGGGCGGACCGTCTTCTCCGACTCAATTTCACTCC	300
QY	643 GATTCCCTCTCTCTCTCTT 662	
Db	301 GATTCCCTCTCTCTCTCTT 320	
RESULT 15		
CG961431/c	954 bp	DNA linear GSS 15-DEC-2003
CG961431		

DEFINITION	MBEKH28TRB mch2 Medicago truncatula genomic clone 73F8, genomic survey sequence.
ACCESSION	CG961431
VERSION	CG961431.1
KEYWORDS	GI:39883077
SOURCE	GSS.
ORGANISM	Medicago truncatula (barrel medic)
REFERENCE	Medicago truncatula Eukaryoto: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eutrosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae; Medicago. 1 (bases 1 to 954)
AUTHORS	Town, C.D., Shetty, J., Koo, H. and Feldblyum, T.F.
TITLE	Sequencing of BAC ends from Medicago truncatula
JOURNAL	Unpublished (2003)
COMMENT	Other GSSs: MBEKH28TRB Contact: Chris Town TIGR
FEATURES	9712 Medical Center Drive, Rockville, MD 20850, USA. Tel: 301-838-3523 Fax: 301-838-0208 Email: cdtown@tigr.org Seq primer: CAGGAACAGCTATGACC Class: BAC ends.
SOURCE	Location/Qualifiers 1..954 /organism="Medicago truncatula" /mol_type="genomic DNA" /culivar="genotype A17" /db_xref="taxon:3880" /clone_id="73F8" /clone_id="mch2" /note="Vector: pBelBAC11; Site_1: HindIII; Site_2: HindIII; Cook, D.R. and Kim, D.J., unpublished"
ORIGIN	
Query Match	8.3%; Score 305.6; DB 10; Length 954;
Best Local Similarity	65.8%; Pred. No. 2.8e-67;
Matches 563;	Conservative 0; Mismatches 239; Indels 54; Gaps 6;
QY	1377 TGTAGGAGGAAGGAGCAAGTATGACCTGTGACACCGGATTTACGTGCAAAATTTGATGAGACTT 1436
DB	856 TGTAGGAGGAAGGAGGAGCAAGCTTAGACCGGATTTTCAAAACAAATTTATATGAGACAC 797
QY	1437 TGAAGAGATCACTCGCGATTATGCTTGAAGCTACTTGGCTTACCGCTTGGATGATTT 1496
DB	796 TTGAAGAGATTAACCTCAGCTGTGTATTAAGAACTTTTAGCTTGCCTCTTATATGATGAC 737
QY	1497 ACGCTGGAAAGACTAATATGTTTAAACGGGTGCGGAATATTTTGTGTCTGTGGAG 1556
DB	736 ATCAGAGCGCGAGGAAGAAAGGTCTCCAAAGGTGCCGCAATCTCTGGGCAAGTTGGAG 677
QY	1557 GAGGTGAGCATGAGCTCTTGTGGGGGTTGACCCCGTGAAGAAATTTATGAATGAGCGT 1616
DB	676 GTGAGGAGGACGACGAATGTCTGGAGTTTCAACGAGAGATTTCAATGATGAGCAT 617
QY	1617 TTTTACGAATGACAGCTGTGAGAGGATATACGTTTAGATA-----CCTTTTAT 1670
DB	616 TCTCGCATATGAAGAGCTGCCGAACAGGTTTCATAGATATATTTGTGATGATTTTAT 557
QY	1671 TTCTTAGCATATATACTTAGGTTTCTCATTTATATGAT----- 1713
DB	556 TTAAATTTCTTCGCCCCCTTAGTTTGTATTAAGAGTTTACATTTATCTCTTCA 497
QY	1714 -----GTGTGTGTAGTTGATCTTTTGTAGTATCCCAAG 1751
DB	496 TTAAATCTTTTATCTCCAAATCTTTTATTTTAAAGTGAACCTTTTGTAGCAACCTAG 437
QY	1752 CAATATTCAGACAGATCATTTTGAAGTTTACGAAGTTGACCTTCTCTTGGGCTCAAG 1811
DB	436 CAATATTCAGCTGAAGTTTGTAGGCGCTATAGGGGTTCACCTTGATTTGTTGACAAAG 377
QY	1812 TTTTATTTGTAGAGCAACCTTTTACAGATCTGATTAAGCAATTTCCAGAACTTCA 1871

Db 376 CTTTGTGTAAGTAAAAAGCCATCTTATTCAGAGATGCTGATTAATTATTCATCAACTTCA 317

Qy 1872 GCAGGCTTAAGGTATAGGCTATGAGATTCCCTGCAGATCTTGTATGATAC-----ACGGAA 1925

Db 316 ACAACACTAAGGTAAACAATATGAGGAATGCTCCCTCTGTTATACTCCCATGAGAGATGA 257

Qy 1926 TAATTGGAGATAGACTCTCGGTCTAGAAAGGGAGCTGTGCACTGCTTATATAGCAAAAGT 1985

Db 256 GAAGAGAGAAGTTCATTTTGATCATTAGAAAGGGGTCTGTGCACTGCTTGTGTGGGAGCT 197

Qy 1986 TGATGAATGCCGTAATGTGCTTGGGCTTGAACAGTGA-GGAATTCACAATATATGGAATCCAG 2044

Db 196 TGATCAATGTGCAATCAATGTTGGGTCTCGATATGACAGTCTCACTTATCGAAACCCAT 137

Qy 2045 CTATGTGAGAGTTGTTTGGAGAAATTCAAATGTGATG-ACAATGATGATCTCCCTGGA 2103

Db 136 CTATTATATGACTTATATATGAGAAACCCAAAGGTATGAAAGTACACTGATCTTCTCCGGA 77

Qy 2104 CTATGCAAAATGTTTGGAAACCTGTTGGACAGGGGTGTCTTTC-CTAGGTTGAGAGAC 2162

Db 76 CTCTGTAATTTGTTGAGACATGGTGTGATGAGAGTGCTTTTCCCTTAGGTTTAAAGATAC 17

Qy 2163 CAAAGTAATAAAAATTT 2178

Db 16 TTAAGAGACAAACTTT 1

Search completed: December 11, 2005, 10:07:53
Job time : 13954.2 secs

QY	2163	CAAAGATAAAAAATT	2178
Db	16	TAAAGAGACAAACTTT	1

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 10, 2005, 08:09:29 ; Search time 2074.73 Seconds
(without alignments)
11779.592 Million cell updates/sec

Title: US-10-600-070B-3

Perfect score: 3667

Sequence: 1 tgcctcgcattcaaggagagat.....ctacacataaaggcctacaa 3667

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_21:*

- 1: geneeqn1980s:*
- 2: geneeqn1990s:*
- 3: geneeqn2000s:*
- 4: geneeqn2001as:*
- 5: geneeqn2001bs:*
- 6: geneeqn2002as:*
- 7: geneeqn2002bs:*
- 8: geneeqn2003as:*
- 9: geneeqn2003bs:*
- 10: geneeqn2003cs:*
- 11: geneeqn2003ds:*
- 12: geneeqn2004as:*
- 13: geneeqn2004bs:*
- 14: geneeqn2005as:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3667	100.0	3667	12	ADJ38130
2	3663.8	99.9	3667	12	ADJ38136
3	1968.4	53.7	2679	13	ADT14901
4	1931	52.7	2637	12	ADJ38210
5	1724	47.0	2406	12	ADJ38129
6	1720.8	46.9	2406	12	ADJ38135
7	1717.6	46.8	2406	12	ADJ38208
8	402.2	11.0	561	12	ADJ38212
9	295.6	8.1	631	12	ADJ38264
10	208.8	5.7	660	12	ADJ38215
11	194	5.3	2283	12	ADJ38206
12	178.6	4.9	607	12	ACN48781
13	175	4.8	439	12	ADP93143
14	161	4.4	545	12	ADJ38255
15	159.4	4.3	552	13	ACN48855
16	158	4.3	608	12	ADJ38217
17	145.6	4.0	527	12	ADJ38214
18	136.2	3.7	647	12	ADJ38265
19	134.6	3.7	652	12	ADJ38266

C	20	132.2	3.6	563	12	ADJ38258	Adj38258
C	21	130.2	3.6	446	12	ADJ38231	Adj38231
C	22	127.6	3.5	479	12	ADJ38230	Adj38230
C	23	125	3.4	537	12	ADJ38223	Adj38223
C	24	123.6	3.3	307	12	ADJ38218	Adj38218
C	25	118.6	3.2	871	12	ADJ38232	Adj38232
C	26	93.6	2.6	491	12	ADJ38254	Adj38254
C	27	92.6	2.5	187	12	ADJ38216	Adj38216
C	28	92	2.5	418	12	ADJ38224	Adj38224
C	29	91.4	2.5	460	12	ADJ38267	Adj38267
C	30	90.4	2.5	420	12	ADJ38256	Adj38256
C	31	89.4	2.4	604	12	ADJ38227	Adj38227
C	32	89.4	2.4	653	12	ADJ38228	Adj38228
C	33	87.2	2.4	480	12	ADJ38225	Adj38225
C	34	86.8	2.4	622	12	ADJ38226	Adj38226
C	35	85.2	2.3	300	12	ADJ38260	Adj38260
C	36	81.8	2.2	360	12	ADJ38259	Adj38259
C	37	80.8	2.2	606	12	ADJ38257	Adj38257
C	38	78	2.1	416	12	ADJ38219	Adj38219
C	39	77.2	2.1	549	12	ADJ38261	Adj38261
C	40	75.8	2.1	540	12	ADJ38233	Adj38233
C	41	66.6	1.8	336	12	ADJ38222	Adj38222
C	42	65.6	1.8	309	12	ADJ38221	Adj38221
C	43	60.2	1.6	295	12	ADJ38213	Adj38213
C	44	59.4	1.6	400	12	ADJ38263	Adj38263
C	45	54	1.5	535	12	ADJ38229	Adj38229

ALIGNMENTS

RESULT 1
ID ADJ38130 standard; DNA; 3667 BP.
XX ADJ38130;
AC
XX
DT 06-MAY-2004 (first entry)
XX
DE Arabidopsis thaliana Arc6-1 genomic DNA SegID2.
XX
XX prokaryotic type; plastid division; Ftn2; ARC6; ARC5; Fzo; plant cell;
XX agronomic; horticultural; crop plant; ornamental plant; woody plant;
XX herbicide target; ds.
XX
OS Arabidopsis thaliana.
XX
XX
XX WO2004001003-A2.
XX
XX
XX 31-DEC-2003.
XX
XX 20-JUN-2003; 2003WO-US019536.
XX
XX 20-JUN-2002; 2002US-0390140P.
XX 09-AUG-2002; 2002US-040242P.
XX 20-JUN-2003; 2003US-0060070.
XX
XX (UNMS) UNIV MICHIGAN STATE.
XX
XX Ostrya Young KW, Vittha S, Kokescharova OA, Gao H;
XX WPI; 2004-082486/08.
XX P-PSDB; ADJ38202.
XX
XX New isolated Ftn2, ARCS and/or Fzo-like nucleic acid sequences, useful
XX for further characterizing plastid division in plant cells, and in
XX varying agronomic and horticultural characteristics of economically
XX important plants.
XX
XX Claim 1; SEQ ID NO 2; 287pp; English.
XX
XX This invention relates to novel prokaryotic type or plastid division and
XX related genes and proteins. In particular, the invention relates to novel

CC Ftn2 (AR6), AR6S and Fzo-like genes and polypeptides. The methods and
CC compositions of the present invention are useful for further
CC characterizing plastid division in plant cells, in order to vary
CC agronomic and horticultural characteristics of economically important
CC plants, such as crop, ornamental and woody plants. They can also be used
CC as herbicide targets. The present sequence is that of a DNA sequence
CC which is related to the invention.

XX Sequence 3667 BP; 983 A; 670 C; 842 G; 1172 T; 0 U; 0 Other;

Query Match 100.0%; Score 3667; DB 12; Length 3667;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 3667; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 TGTTCGATTAGAGAGATACAAATTATAGCAATTTGCTTGATTTCAAGAAATTTG 60
DB 1 TGTTCGATTAGAGAGATACAAATTATAGCAATTTGCTTGATTTCAAGAAATTTG 60
QY 61 CTGGGCTATAGGATTCATTGGCTCTGTTGCTTTTACATTTACATGTCATTAATGTTGG 120
DB 61 CTGGGCTATAGGATTCATTGGCTCTGTTGCTTTTACATTTACATGTCATTAATGTTGG 120
QY 121 AATTTTACACATTTGATGTTGATTTAAGAAAGAGAGAAATGATGGGGTTTTGTTGG 180
DB 121 AATTTTACACATTTGATGTTGATTTAAGAAAGAGAGAAATGATGGGGTTTTGTTGG 180
QY 181 TTTAACTTTAAAGTACAGAAATTAAGTCATTTGTTACTGTTCTATATGTTTAA 240
DB 181 TTTAACTTTAAAGTACAGAAATTAAGTCATTTGTTACTGTTCTATATGTTTAA 240
QY 241 AATGAAGCACTCCAACGGTCTTAGGTGATGATTTATTTAGACATTTAAACATCAT 300
DB 241 AATGAAGCACTCCAACGGTCTTAGGTGATGATTTATTTAGACATTTAAACATCAT 300
QY 301 AAAGTCGTTGGGACTGTATACATCATAGATTGTTTTTATTTTTCAGTGTGAT 360
DB 301 AAAGTCGTTGGGACTGTATACATCATAGATTGTTTTTATTTTTCAGTGTGAT 360
QY 361 GTTTTTGATTTAACTTAACTTAACTCAAAATCAAAATTCATTAACCTTAGACGACAA 420
DB 361 GTTTTTGATTTAACTTAACTTAACTCAAAATCAAAATTCATTAACCTTAGACGACAA 420
QY 421 CAGTCTCTTCAATATGTAAACAGAAAGTTTGTATGAGCTTAAAGACATCTCC 480
DB 421 CAGTCTCTTCAATATGTAAACAGAAAGTTTGTATGAGCTTAAAGACATCTCC 480
QY 481 ATGGAAGCTGAGTCAAGTGGGCAATGGTCTCTCCCATCCATTCOAATTAAGCGATTACA 540
DB 481 ATGGAAGCTGAGTCAAGTGGGCAATGGTCTCTCCCATCCATTCOAATTAAGCGATTACA 540
QY 541 CCGGCGACGACAAAGCTCGAGTACGACAAACCTCTACAACTATCTGCTCGCCAGC 600
DB 541 CCGGCGACGACAAAGCTCGAGTACGACAAACCTCTACAACTATCTGCTCGCCAGC 600
QY 601 AAATGGGCGGACCGTCTTCTCTCCGACTTCAATTTCACTCCGATCTCTCTCTCC 660
DB 601 AAATGGGCGGACCGTCTTCTCTCCGACTTCAATTTCACTCCGATCTCTCTCTCC 660
QY 661 TTGCGCACCGGACACGCGGACGCTGCGTCTGCGACCACTATTAATGATGTTGCC 720
DB 661 TTGCGCACCGGACACGCGGACGCTGCGTCTGCGACCACTATTAATGATGTTGCC 720
QY 721 GAAGCGCACGTCGCCATCCCATTTGATTTCTACAGATTTAGAGCTCAAAACATTTTC 780
DB 721 GAAGCGCACGTCGCCATCCCATTTGATTTCTACAGATTTAGAGCTCAAAACATTTTC 780
QY 781 TTTAACCGATGGAATGAAAGAGCATTTGAGTATGAGGTTTGAACCCGCGCAATTCGGT 840
DB 781 TTTAACCGATGGAATGAAAGAGCATTTGAGTATGAGGTTTGAACCCGCGCAATTCGGT 840
QY 841 TTGAGCGGACGCTTTAATGACCGGAGAGACGATTTTCAAGCTGTGGAAACCTGTG 900
DB 841 TTGAGCGGACGCTTTAATGACCGGAGAGACGATTTTCAAGCTGTGGAAACCTGTG 900
```

```
QY 901 TCTAATCCTCGTCTAGAGAGATACAAATGAAGTCTTCTGATGATGAAGAGCTACA 960
DB 901 TCTAATCCTCGTCTAGAGAGATACAAATGAAGTCTTCTGATGATGAAGAGAGCTACA 960
QY 961 GTCATCTGATGTTCTTGGGATTAAGTAAATTTCCATTTCCGAAATATATAAGTTCTTC 1020
DB 961 GTCATCTGATGTTCTTGGGATTAAGTAAATTTCCATTTCCGAAATATATAAGTTCTTC 1020
QY 1021 GTTTAATTTTCATGAATTTGGATTAAGAGAGAACTTTATCTAGTGAAGGTTCTGGGGG 1080
DB 1021 GTTTAATTTTCATGAATTTGGATTAAGAGAGAACTTTATCTAGTGAAGGTTCTGGGGG 1080
QY 1081 TCTCTGTGATTTGCAAGAGGTGTGAGACTGAGATAGTTCTTCGGTTGTGAGGCTCT 1140
DB 1081 TCTCTGTGATTTGCAAGAGGTGTGAGACTGAGATAGTTCTTCGGTTGTGAGGCTCT 1140
QY 1141 GCTTAAGAGAGGTGTGCTTAAGTCTTAAAGCAAGATGTTGTTAATGAGGCTTTC 1200
DB 1141 GCTTAAGAGAGGTGTGCTTAAGTCTTAAAGCAAGATGTTGTTAATGAGGCTTTC 1200
QY 1201 GTTTCGATGTCGAGGATGCTATGCAATTCACCTGATTTTATTAATCTGGTTA 1260
DB 1201 GTTTCGATGTCGAGGATGCTATGCAATTCACCTGATTTTATTAATCTGGTTA 1260
QY 1261 TGAATTTGTTGAGGAACCTTTGAAGCTTTTACAGGTAAGTTGACTTGTGTAATTTG 1320
DB 1261 TGAATTTGTTGAGGAACCTTTGAAGCTTTTACAGGTAAGTTGACTTGTGTAATTTG 1320
QY 1321 ACAGCGTGTGCTTTAATAAGACTTCTTGATTTGAATCTTTGTATGATGCTTGTTGA 1380
DB 1321 ACAGCGTGTGCTTTAATAAGACTTCTTGATTTGAATCTTTGTATGATGCTTGTTGA 1380
QY 1381 GAGAGAGAGAGCAAGTACCTTGCAACCGAATTAACGTGCAAAATGATAGACTTTGA 1440
DB 1381 GAGAGAGAGAGCAAGTACCTTGCAACCGAATTAACGTGCAAAATGATAGACTTTGA 1440
QY 1441 AGAGATCACTCCGCTTATGTTGAGACTTGTGAGCTTGTGCTTCCGCTTGTGATTAAGC 1500
DB 1441 AGAGATCACTCCGCTTATGTTGAGACTTGTGAGCTTGTGCTTCCGCTTGTGATTAAGC 1500
QY 1501 TCGCAAAAAGCTAATATGTTTAAAGCGGTGCGGAATATTTTGTGTGTGAGAGAG 1560
DB 1501 TCGCAAAAAGCTAATATGTTTAAAGCGGTGCGGAATATTTTGTGTGTGAGAGAG 1560
QY 1561 TGGAGCATCAGCTTGTGTGGGGTTTGACCCGAGAGAACTTTATGAAATGAGCGTTT 1620
DB 1561 TGGAGCATCAGCTTGTGTGGGGTTTGACCCGAGAGAACTTTATGAAATGAGCGTTT 1620
QY 1621 ACGAATGACAGCTGTGAGAGAGTATACAGTTTATGATTAACCTTTTATTTCTTTAGCA 1680
DB 1621 ACGAATGACAGCTGTGAGAGAGTATACAGTTTATGATTAACCTTTTATTTCTTTAGCA 1680
QY 1681 TGAATATACTTTAGGTTTCAATTTTAAATGATGTTGTTGAGTGAATCTTTTGTGA 1740
DB 1681 TGAATATACTTTAGGTTTCAATTTTAAATGATGTTGTTGAGTGAATCTTTTGTGA 1740
QY 1741 GCTACCCCAAGCAATATTCAGAGAGTCAATTTGAAGTTTACGAAGTTGACTTGTCTT 1800
DB 1741 GCTACCCCAAGCAATATTCAGAGAGTCAATTTGAAGTTTACGAAGTTGACTTGTCTT 1800
QY 1801 GTGGCTCAAGCTTTATTTGTTAAGAGCCACACTTTTACAGATGCTGATTAAGCAATTC 1860
DB 1801 GTGGCTCAAGCTTTATTTGTTAAGAGCCACACTTTTACAGATGCTGATTAAGCAATTC 1860
QY 1861 CAGCAACTTTACAGAGCTTAAGTATATGCTATGAGATTTCCGCAAGTTGTATGATACA 1920
DB 1861 CAGCAACTTTACAGAGCTTAAGTATATGCTATGAGATTTCCGCAAGTTGTATGATACA 1920
QY 1921 CCGAATAATTTGGAGATAGACTTCGGTCTTGAAGAGGGACTCTGTGCACTGCTTATAGGC 1980
DB 1921 CCGAATAATTTGGAGATAGACTTCGGTCTTGAAGAGGGACTCTGTGCACTGCTTATAGGC 1980
```


XX WPI: 2004-082486/08.
DR P-PSDB; ADJ38203.
XX
XX New isolated Ftn2, ARCS and/or Fzo-like nucleic acid sequences, useful
PT for further characterizing plastid division in plant cells, and in
PT varying agronomic and horticultural characteristics of economically
PT important plants.
PS
PS Claim 1; SEQ ID NO 10; 287bp; English.
XX
XX This invention relates to novel prokaryotic type or plastid division and
CC related genes and proteins. In particular, the invention relates to novel
CC Ftn2 (ARCS), ARCS and Fzo-like genes and polypeptides. The methods and
CC compositions of the present invention are useful for further
CC characterizing plastid division in plant cells, in order to vary
CC agronomic and horticultural characteristics of economically important
CC plants, such as crop, ornamental and woody plants. They can also be used
CC as herbicide targets. The present sequence is that of a DNA sequence
CC which is related to the invention.
XX
XX

Sequence 3667 BP; 982 A; 669 C; 843 G; 1173 T; 0 U; 0 Other;

Query Match 99.9%; Score 3663.8; DB 12; Length 3667;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 3665; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TGTTCGATTAAGAGAAATCAATTAATAGCAATTTGCTTGATTTCAACAAGATTTTG 60
DB 1 TGTTCGATTAAGAGAAATCAATTAATAGCAATTTGCTTGATTTCAACAAGATTTTG 60
QY 61 CTGGGATATAGGATTCATTGAGCTGTGTTGCTTTTACATTTACATGTCATAATAGTTTCG 120
DB 61 CTGGGATATAGGATTCATTGAGCTGTGTTGCTTTTACATTTACATGTCATAATAGTTTCG 120
QY 121 AATTTTACATTTTCAAGTGTGATGTTAAGAAAAGAGAGGATTTGATGGGTTTGTGGG 180
DB 121 AATTTTACATTTTCAAGTGTGATGTTAAGAAAAGAGAGGATTTGATGGGTTTGTGGG 180
QY 181 TTTTAACTTTAAGTATGTCAGAAATTAAGTCAATTTGCTTACTGCTATATATGTATA 240
DB 181 TTTTAACTTTAAGTATGTCAGAAATTAAGTCAATTTGCTTACTGCTATATATGTATA 240
QY 241 AATGAAGGCACTCCACCGGTTCTTAGTGAGATAGATTATTAGACGATTTAACATCAT 300
DB 241 AATGAAGGCACTCCACCGGTTCTTAGTGAGATAGATTATTAGACGATTTAACATCAT 300
QY 301 AAAGTCCGTGGGACGTGTAACATCATAGATGTTTTTATTTTTCAGTACCTGTGAT 360
DB 301 AAAGTCCGTGGGACGTGTAACATCATAGATGTTTTTATTTTTCAGTACCTGTGAT 360
QY 361 GTTTTTGTATTAACCTTACTACTCAAAATCAAAATTCATTAACCCCTAGACGACAAA 420
DB 361 GTTTTTGTATTAACCTTACTACTCAAAATTCATTAACCCCTAGACGACAAA 420
QY 421 CAGTCTCTTCAATATATGTAAGACAAAGTTTGTAGTAGCTTAAGAAAGACACTCC 480
DB 421 CAGTCTCTTCAATATATGTAAGACAAAGTTTGTAGTAGCTTAAGAAAGACACTCC 480
QY 481 ATGGAAGCTCTAGTCACTCGGCAATGTGCTCTCCCAATTCCAATTAGCGGATTAACA 540
DB 481 ATGGAAGCTCTAGTCACTCGGCAATGTGCTCTCCCAATTCCAATTAGCGGATTAACA 540
QY 541 CCGGCGACGACAAAGCTCGAGGTAGCCACAAACCTCTACAACTATCTGCTCCGACGC 600
DB 541 CCGGCGACGACAAAGCTCGAGGTAGCCACAAACCTCTACAACTATCTGCTCCGACGC 600
QY 601 AATGGGCGGACCGTCTCTCCGACTTCAATTTCACTCCGATTCCTCTCTCTCC 660
DB 601 AATGGGCGGACCGTCTCTCTCCGACTTCAATTTCACTCCGATTCCTCTCTCTCC 660
QY 661 TTGCGCACCGCACCAACCGGCACTCTGCTCTCTGCGACCATTTATGATGTCC 720
DB 661 TTGCGCACCGCACCAACCGGCACTCTGCTCTCTGCGACCATTTATGATGTCC 720

DB 661 TTGCGCACCGCACCAACCGGCACTCTGCTCTCTGCGACCATTTATGATGTCC 720
QY 721 GAAGCCACGTCCTCCATCCCATTTGATTTTACAGGATATTAGAGCTCAACATTTTC 780
DB 721 GAAGCCACGTCCTCCATCCCATTTGATTTTACAGGATATTAGAGCTCAACATTTTC 780
QY 781 TTAACCGATGGAATCAGAAAGCATTTGAAAGCTTAGGGTTTCGAAACCGCCGCAATTGGT 840
DB 781 TTAACCGATGGAATCAGAAAGCATTTGAAAGCTTAGGGTTTCGAAACCGCCGCAATTGGT 840
QY 841 TTGAGCGACGCGTTTATTCAGCCGAGACAGATTCTTCAAGCTGCTTGGAACCTG 900
DB 841 TTGAGCGACGCGTTTATTCAGCCGAGACAGATTCTTCAAGCTGCTTGGAACCTG 900
QY 901 TCTATCTCGGCTTGAAGAGAGTACATGAAGGCTTCTTGATGATGAAGAGCTACA 960
DB 901 TCTATCTCGGCTTGAAGAGAGTACATGAAGGCTTCTTGATGATGAAGAGCTACA 960
QY 961 GTCATCATGATGTTCTTGAGATAGGTAATTTGATTTGCAATTCGAAATATAAGTTCTTC 1020
DB 961 GTCATCATGATGTTCTTGAGATAGGTAATTTGATTTGCAATTCGAAATATAAGTTCTTC 1020
QY 1021 GTTTTATTTCAATTAATGATTAAGAAAGAACTTTATCTAGTGAAGGTTCTGAGGCG 1080
DB 1021 GTTTTATTTCAATTAATGATTAAGAAAGAACTTTATCTAGTGAAGGTTCTGAGGCG 1080
QY 1081 TCTCTGTGTATTTGCAAGAGGTGCTGAGACTGATAGTATTTCTCGGTTGGTGAAGCTCT 1140
DB 1081 TCTCTGTGTATTTGCAAGAGGTGCTGAGACTGATAGTATTTCTCGGTTGGTGAAGCTCT 1140
QY 1141 GCTTAAGAGAGGTGCTTAAGTCTGTTAAGCAAGATGCTGTTTATGATGAGCGCTTGC 1200
DB 1141 GCTTAAGAGAGGTGCTTAAGTCTGTTAAGCAAGATGCTGTTTATGATGAGCGCTTGC 1200
QY 1201 GTTTCGATGCTCGAGGAGATGCTATGAGATGCACTGATTTTATCTGAGTTA 1260
DB 1201 GTTTCGATGCTCGAGGAGATGCTATGAGATGCACTGATTTTATCTGAGTTA 1260
QY 1261 TGAATTTGTGAGGAACTTTGAAGCTTTTACAGATAGTTGACTTTCGTTGATTTTG 1320
DB 1261 TGAATTTGTGAGGAACTTTGAAGCTTTTACAGATAGTTGACTTTCGTTGATTTTG 1320
QY 1321 ACGAGCGTGGCTTTATTAAGAACTTTCTGATTTGATTAAGTGTGATGCTGTGCTGA 1380
DB 1321 ACGAGCGTGGCTTTATTAAGAACTTTCTGATTTGATTAAGTGTGATGCTGTGCTGA 1380
QY 1381 GGAAGAGAGGCAAGTACCTGCAACGGAATTAACGTCGCAAAATGATAGACTTTGGA 1440
DB 1381 GGAAGAGAGGCAAGTACCTGCAACGGAATTAACGTCGCAAAATGATAGACTTTGGA 1440
QY 1441 AGAGATCACTCCGCTTATGCTTGGAGCTACTTGGCTTACCGCTTGATGATTAAGC 1500
DB 1441 AGAGATCACTCCGCTTATGCTTGGAGCTACTTGGCTTACCGCTTGATGATTAAGC 1500
QY 1501 TGGCAAAAGACTAATGTTTAAAGCGGTGCGGAATATTTTGTGCTGTGGAGAGG 1560
DB 1501 TGGCAAAAGACTAATGTTTAAAGCGGTGCGGAATATTTTGTGCTGTGGAGAGG 1560
QY 1561 TGGAGCATGCTCTTGTGGGGGTTTGACCCGGAAGATTTATGATGAATGAGGGGTTTTT 1620
DB 1561 TGGAGCATGCTCTTGTGGGGGTTTGACCCGGAAGATTTATGATGAATGAGGGGTTTTT 1620
QY 1621 ACGATGACAGCTCTGAGCAGATACAGTTTATAGATACCTTTTATTTTCTTTAGCA 1680
DB 1621 ACGATGACAGCTCTGAGCAGATACAGTTTATAGATACCTTTTATTTTCTTTAGCA 1680
QY 1681 TGATATACTTTAGGTTTCTCATTTTATATGATGTTGTGATAGTTGATCTTTTGTGA 1740
DB 1681 TGATATACTTTAGGTTTCTCATTTTATATGATGTTGTGATAGTTGATCTTTTGTGA 1740
QY 1741 GCTACCCCAACCATATTTCCAGAGAGCTTTGAAGTTTACGAAGTTGACCTTGCTCTT 1800
DB 1741 GCTACCCCAACCATATTTCCAGAGAGCTTTGAAGTTTACGAAGTTGACCTTGCTCTT 1800

[illegible]

QY	2881	CTGTAGACAGAAATGTAAATTTCTACTCTCAACATTTCTGTTTGAATAAGTAGATTAAG	2940
Db	2881	CTGTAGACAGAAATGTAAATTTCTACTCTCAACATTTCTGTTTGAATAAGTAGATTAAG	2940
QY	2881	CTGTAGACAGAAATGTAAATTTCTACTCTCAACATTTCTGTTTGAATAAGTAGATTAAG	2940
Db	2881	CTGTAGACAGAAATGTAAATTTCTACTCTCAACATTTCTGTTTGAATAAGTAGATTAAG	2940
QY	2941	AGATTCCTTAGTGTGGCTTTGTGCCAATTTTCTTTTCTTGATTTTCTTTTGCATTT	3000
Db	2941	AGATTCCTTAGTGTGGCTTTGTGCCAATTTTCTTTTCTTGATTTTCTTTTGCATTT	3000
QY	3001	AGGTCAGTCAGAGCTGACGATTTCAAGACATTTCCAGAAATGGATGCTAGAGATTCGACA	3060
Db	3001	AGGTCAGTCAGAGCTGACGATTTCAAGACATTTCCAGAAATGGATGCTAGAGATTCGACA	3060
QY	3061	GAATTTATGTATCCAAAGTGGCAGAAAGATTAAAGTCCTGGCTTTTGGGCTGATCCGCAT	3120
Db	3061	GAATTTATGTATCCAAAGTGGCAGAAAGATTAAAGTCCTGGCTTTTGGGCTGATCCGCAT	3120
QY	3121	AGAAATGTTTACAGAGGTGAGGGAAATTAATCTACAAATTCATCAATTTGTGTGAATCTGT	3180
Db	3121	AGAAATGTTTACAGAGGTGAGGGAAATTAATCTACAAATTCATCAATTTGTGTGAATCTGT	3180
QY	3181	TGACATGATTTATAGTCCTGGCTTTGTTGATTTCTGTATTATTAAGGTTTGGATGGGC	3240
Db	3181	TGACATGATTTATAGTCCTGGCTTTGTTGATTTCTGTATTATTAAGGTTTGGATGGGC	3240
QY	3241	GAATGCTGAAGATTTGGAACCTGACAGACACAGCTGAAATCGGCGAGCTGGGTGGTTATG	3300
Db	3241	GAATGCTGAAGATTTGGAACCTGACAGACACAGCTGAAATCGGCGAGCTGGGTGGTTATG	3300
QY	3301	ATTATTAACACTGTTGAAACTATCTGTTGACAGTGTGACAGTCTCGACAGATGSAACCCGTG	3360
Db	3301	ATTATTAACACTGTTGAAACTATCTGTTGACAGTGTGACAGTCTCGACAGATGSAACCCGTG	3360
QY	3361	CTCTGTGTGAAGCAACTCTGGAGGAGTCTGCTTGTCTATCTGATTTGGTTTCATCAGAAA	3420
Db	3361	CTCTGTGTGAAGCAACTCTGGAGGAGTCTGCTTGTCTATCTGATTTGGTTTCATCAGAAA	3420
QY	3421	ACAAATGCTACTGATGTGAGAACTTACACAAACAAAGATTCGAAGTTTCTGGTCCAAAGTCAG	3480
Db	3481	GGTGGAAATACACTGGAAGGCTCTGTTCTTGACATCATATTAATTAATGATGATGCTCT	3540
QY	3541	GAGCTTGCAGAAATCTCTTTGTTCTGTAAATTCCTCTCTAAGTTAGTGTATTAATAATGA	3600
Db	3541	GAGCTTGCAGAAATCTCTTTGTTCTGTAAATTCCTCTCTAAGTTAGTGTATTAATAATGA	3600
QY	3601	ACACAAAAAAATTAAGCTTCTTGCCACACCCCTTTCTTGATCTAAACCTAATACATAAGG	3660
Db	3601	ACACAAAAAAATTAAGCTTCTTGCCACACCCCTTTCTTGATCTAAACCTAATACATAAGG	3660
QY	3661	GCTACAA 3667	
Db	3661	GCTACAA 3667	
RESULT 3			
ADT14901			
ID	ADT14901	standard; cDNA; 2679 BP.	
XX	ADT14901;		
XX	13-JAN-2005	(first entry)	
DE	Plant cDNA, Seq ID 227.		
KW	Plant; ss; gene; transgenic; cold tolerance; growth rate;		
KW	drought tolerance; disease resistance; galactomannan production;		
KW	plant growth regulator; heat tolerance; herbicide tolerance;		
KW	lignin production; extreme osmotic condition tolerance;		
KW	pathogens resistance; pest resistance; yield improvement; seed oil yield;		
KW	seed protein yield.		

XX Viridiplantae.
OS
XX US2004216190-A1.
XX
XX 28-OCT-2004.
XX
XX 18-DEC-2003; 2003US-00739930.
XX
XX 28-APR-2003; 2003US-00424599.
XX 28-APR-2003; 2003US-00425115.
XX
XX (KOVA/) KOVALIC D K.
XX
XX Kovalic DK;
XX WPI, 2004-757369/74.
XX
XX New recombinant DNA constructs useful in the field of biochemistry and
PT genetics, and in particular for producing transgenic plants with improved
PT biological characteristics.
XX
XX Claim 1; SEQ ID NO 227; 14pp; English.
XX
XX The invention relates a recombinant DNA construct comprising a
CC polynucleotide having any of 5544 nucleotide sequences (CDNs SEQ ID NO:
CC 1-5544) and encoding a polypeptide with any of 5544 amino acid sequences
CC (SEQ ID NO: 5545-11088). The CDNs and proteins are from corn, soybean,
CC Arabidopsis, wheat and rape but the specification does not indicate which
CC sequences is derived from which organism. Also included is a method of
CC producing a plant having an improved property, comprising transforming a
CC plant with a recombinant DNA construct comprising a promoter region
CC functional in a plant cell operably joined to a polynucleotide encoding a
CC polypeptide associated with the property, and growing the transformed
CC plant. The property is selected from improving plant cold tolerance, for
CC manipulating growth rate in plant cells by modification of the cell cycle
CC pathway, for improving plant drought tolerance, for providing increased
CC resistance to plant diseases, for galactomanan production, for production
CC of plant growth regulators, for improving plant heat tolerance, for
CC improving plant tolerance to herbicides, for increasing the rate of
CC homologous recombination in plants, for lignin production, for improving
CC plant tolerance to extreme osmotic conditions, for improving plant
CC tolerance to pathogens or pests, for yield improvement by modification of
CC photosynthesis, for modifying seed oil yield and/or content, for
CC modifying seed protein yield and/or content, for yield improvement by
CC modification of carbohydrate, nitrogen or phosphorus use and/or uptake
CC and for yield improvement by providing improved plant growth and
CC development under at least one stress condition. The polynucleotide may
CC also encode a plant transcription factor. The methods and compositions of
CC the present invention are useful in the field of biochemistry and
CC genetics, in particular for producing transgenic plants with improved
CC biological characteristics such as increased yield, improved nitrogen
CC flow, increasing plant tolerance to cold or heat, improving plant
CC tolerance to extreme osmotic and drought conditions, and improving plant
CC tolerance to plant pests or pathogens. They can also be used in physical
CC arrays of molecules, plant breeding markers, computer-based storage and
CC analysis systems. The present sequence is one of the 5544 plant cDNA
CC sequences of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20040216190.
XX
XX Sequence 2679 BP; 712 A; 545 C; 653 G; 769 T; 0 U; 0 Other;
SQ
Query Match 53.7%; Score 1968.4; DB 13; Length 2679;
Best Local Similarity 80.5%; Pred. No. 0;
Matches 2657; Conservative 0; Mismatches 11; Indels 632; Gaps 5;

Db
62 TTCAATATGTAAACAAACAAAGTTTGTAGAGCTTAAAAAGACATCCATGGAAG 121
OY
488 CTCTGAGTCAGCGGATTTGCTCTCCCATTCCTAATATGACGATTAACACCGCGA 547
Db
122 CTCTGAGTCAGCGGATTTGCTCTCCCATTCCTAATATGACGATTAACACCGCGA 181
OY
548 CGACAAAGCTCCGACGTAAGCAACAACTCTCAACATATCTGCTCCGACGAAATGAG 607
Db
182 CGACAAAGCTCCGACGTAAGCAACAACTCTCAACATATCTGCTCCGACGAAATGAG 241
OY
608 CCGACCGCTTTCTCTCCGACTTCAATTTACCTCCGATTCCTCTCTCTCTCTGCGCA 667
Db
242 CCGACCGCTTTCTCTCCGACTTCAATTTACCTCCGATTCCTCTCTCTCTCTGCGCA 301
OY
668 CCGCACACACACCGGCACTCTGCTCTCTGACACCATCTATTTAGTCGCCGAAGCC 727
Db
302 CCGCACACACACCGGCACTCTGCTCTCTGACACCATCTATTTAGTCGCCGAAGCC 361
OY
728 ACGTCCCATCCCATTTGATTTTACCAAGGATTTAGAGCTCAACACATTTCTTAACG 787
Db
362 ACGTCCCATCCCATTTGATTTTACCAAGGATTTAGAGCTCAACACATTTCTTAACG 421
OY
788 ATGAATCAGAAAGCATTTGAAAGCTAGGTTTGGAAACCGCCGCAATTTGCTTACGC 847
Db
422 ATGAATCAGAAAGCATTTGAAAGCTAGGTTTGGAAACCGCCGCAATTTGCTTACGC 481
OY
848 ACGACGCTTTAATCAGCGGACAGACAGATTTCTCAAGCTGCTGCAAACTCTGCTATC 907
Db
482 ACGACGCTTTAATCAGCGGACAGACAGATTTCTCAAGCTGCTGCAAACTCTGCTATC 541
OY
908 CTCGGTCTAGAAAGAGTACATGAAGGCTCTTGTATGATGAAGAGCTACAGTCATCA 967
Db
542 CTCGGTCTAGAAAGAGTACATGAAGGCTCTTGTATGATGAAGAGCTACAGTCATCA 601
OY
968 CTGATTTCTCTTGGGATTAAGGTAATTCGATTTCCGAAATATAAGTTCTCTGTTTA 1027
Db
602 CTGATTTCTCTTGGGATTAAGGTAATTCGATTTCCGAAATATAAGTTCTCTGTTTA 618
OY
1028 TTTATGAAATGTGATTAAGAAAGAACTTTATCTAGTGAAGTTCTTGAGGCTCTCTGT 1087
Db
619 TTTATGAAATGTGATTAAGAAAGAACTTTATCTAGTGAAGTTCTTGAGGCTCTCTGT 639
OY
1088 GTATTTGCAAGAGGTGTGAGACTGAGATGATTTCTTGAGGCTGTGAGGCTCTGTTAAG 1147
Db
640 GTATTTGCAAGAGGTGTGAGACTGAGATGATTTCTTGAGGCTGTGAGGCTCTGTTAAG 699
OY
1148 GAGAGGTTGCTTAAGTGTGTTAAGCAAGATGTTTATGTTATGAGGCTGCTTCTC 1207
Db
700 GAGAGGTTGCTTAAGTGTGTTAAGCAAGATGTTTATGTTATGAGGCTGCTTCTC 759
OY
1208 GATGCTCGAAGGAGTGTGAGCTGATTCATCTGATTTTATCTGTTATGAGTTT 1267
Db
760 GATGCTCGAAGGAGTGTGAGCTGATTCATCTGATTCATCTGTTATGAGTTT 819
OY
1268 GTTGAGGAAGCTTTGAAGCTTTTACAGAGTGTGAGCTGTTGTTGTAATTTGACGAGCG 1327
Db
820 GTTGAGGAAGCTTTGAAGCTTTTACAGAGTGTGAGCTGTTGTTGTAATTTGACGAGCG 844
OY
1328 TTGGCTTTATGAAGACTTTCTTGATTTGATTAATTTGATGAGCTTGTGTGAGAGAA 1387
Db
845 TTGGCTTTATGAAGACTTTCTTGATTTGATTAATTTGATGAGCTTGTGTGAGAGAA 852
OY
1388 GGAGCAAGTAGCTTTGACCGGATTTTACGTCACAAATTTGATGAGACTTTGGAAGATC 1447
Db
853 GGAGCAAGTAGCTTTGACCGGATTTTACGTCACAAATTTGATGAGACTTTGGAAGATC 912
OY
1448 ACTCGGCTTATGCTTGTGAGCTTACCTTGCTTACCGCTTGATGATTTACGCTCGAAA 1507
Db
913 ACTCGGCTTATGCTTGTGAGCTTACCTTGCTTACCGCTTGATGATTTACGCTCGAAA 972
OY
1508 AGACTGAATGTTTAAAGCGGTGCGGAATTTTGTGTGCTGTTGAGAGAGGTGAGCA 1567

Db 973 AGACTAAATGTTTAAAGCGTGTGCGGAATATTTTGCTGTTGAGAGAGCTGAGCA 1032
Qy 1568 TCAGCTCTGTGGGGGTTTGAACCCGTGAGAGATTATGATGAGCCGTTTTTACGAATG 1627
Db 1023 TCAGCTCTGTGGGGGTTTGAACCCGTGAGAGATTATGATGAGCCGTTTTTACGAATG 1092
Qy 1628 ACAGCTGCTGAGCAGGTATACAGTTTAAATCTTTTAAATTTCTTTAGCATATATA 1687
Db 1093 ACAGCTGCTGAGC----- 1105
Qy 1688 ACTTAGGTTTCTCATTTTAAATGATATGTTGTGTGAGTTGATCTTTTGTAGTACC 1747
Db 1106 -----AGTTGATCTTTTGTAGTACC 1129
Qy 1748 CAAGCAATATTCAGACAGATCATTTGAAGTTTACGAATGTCATTCTCTGTGGCTC 1807
Db 1130 CAAGCAATATTCAGACAGATCATTTGAAGTTTACGAATGTCATTCTCTGTGGCTC 1189
Qy 1808 AAGCTTTATTTGTTAAGAACCAACCTTTTACAGAGTGTGATTAAGCAATTCAGCAAC 1867
Db 1190 AAGCTTTATTTGTTAAGAACCAACCTTTTACAGAGTGTGATTAAGCAATTCAGCAAC 1249
Qy 1868 TTCACAGAGCTAAGGTATAGGTATGAGATTCCTGCGATGTGATATACAGCAATA 1927
Db 1250 TTCACAGAGCTAAGGTATAGGTATGAGATTCCTGCGATGTGATATACAGCAATA 1309
Qy 1928 ATTGGAGATAGACTTCGGTCTAGAAAGGGACTCTGTGCACTGCTTATAGGCAAAAGTTG 1987
Db 1310 ATTGGAGATAGACTTCGGTCTAGAAAGGGACTCTGTGCACTGCTTATAGGCAAAAGTTG 1369
Qy 1988 ATGAATGCCGATGTGTTGGGCTTAAACAGTGAAGATTCACAAATATAGAAATCCAGCTA 2047
Db 1370 ATGAATGCCGATGTGTTGGGCTTAAACAGTGAAGATTCACAAATATAGAAATCCAGCTA 1429
Qy 2048 TTGGGAGGTTGTTTGGAGAAATTCAAATCGATGAAATGAAATGATCCCTGCAATAT 2107
Db 1430 TTGGGAGGTTGTTTGGAGAAATTCAAATCGATGAAATGAAATGATCCCTGCAATAT 1489
Qy 2108 GCAAAATGTTGGAACCTGTTGAGGAGGTTGCTTCTAGAGTTACAGAACCAAAAG 2167
Db 1490 GCAAAATGTTGGAACCTGTTGAGGAGGTTGCTTCTAGAGTTACAGAACCAAAAG 1549
Qy 2168 ATAAAAATTTAAACTCGGGGACTATATGATGATCCTATGTTTGAAGTTACTTGGAAA 2227
Db 1550 ATAAAAATTTAAACTCGGGGACTATATGATGATCCTATGTTTGAAGTTACTTGGAAA 1609
Qy 2228 GAGTGAAGGTAGTTCAAGGTTCTCCTTATAGCTGCTGCTGCAATATGCAAGATTTGAG 2287
Db 1610 GAGTGAAGGTAGTTCAAGGTTCTCCTTATAGCTGCTGCTGCAATATGCAAGATTTGAG 1669
Qy 2288 CCGAGCATGTGAAGCTAGTCTATGACAGGACTGACAGAAAGTTTCTCCGCTATA 2347
Db 1670 CCGAGCATGTGAAGCTAGTCTATGACAGGACTGACAGAAAGTTTCTCCGCTATA 1729
Qy 2348 CAGATTAAGAACTCGGCTGAACCCCAAGATGTGCAAGAGACAGTGTATAGTAAATCCGTG 2407
Db 1730 CAGATTAAGAACTCGGCTGAACCCCAAGATGTGCAAGAGACAGTGTATAGTAAATCCGTG 1789
Qy 2408 TTGGTAAACAATGTAGGCGGTATGATGTGAGCTGTGTCTTATATGCAAACTGTAAAGAC 2467
Db 1790 TTGGTAAACAATGTAGGCGGTATGATGTGAGCTGTGTCTTATATGCAAACTGTAAAGAC 1849
Qy 2468 CCTCTGAATACTTTGAATACTATGATTAATGCAATTCGAGCTGGGGTCTCAGAGAGTACG 2527
Db 1850 CCTCTGAATACTTTGAATACTATGATTAATGCAATTCGAGCTGGGGTCTCAGAGAGTACG 1909
Qy 2528 TTGATGAATACTAGTGTGAATAATGTCGTTGCTGATATGTTAAAGAGGCAAGTGTGAAGA 2587
Db 1910 TTGATGAATACTAGTGTGAATAATGTCGTTGCTGATATGTTAAAGAGGCAAGTGTGAAGA 1969
Qy 2588 TCCATAGCTGCTGAGTGTGCAATTTGCACTGATTTCACTGTTTCAAGCAAGATTTTCTTA 2647
Db 1970 TCCATAGCTGCTGAGTGTGCAATTTGCACTGATTTCACTGTTTCAAGCAAGATTTTCTTA 2029

Qy 2648 AAAGAGCTCATCTTTTCAACGCAAGATATAGTGTCTTCTATAGAAATCTGATGTGCTTA 2707
Db 2030 AAAGAGCTCATCTTTTCAACGCAAGATATAGTGTCTTCTATAGAAATCTGATGTGCTTA 2089
Qy 2708 CCATAGGTATATTAATGATGCAATTTTCAATATCTGCAATGTCTCAAAAATATGCTTGT 2767
Db 2090 CCA----- 2092
Qy 2768 TTGTGAGCTAAGAACATATGTTCCACTAATATGATGCCAAAAGTTGTACCAAGATTA 2827
Db 2093 ----- 2092
Qy 2828 ACAAGTCTGAGTAAATTTCACTAATATATGCTGTAATTTTGTATCAAACTGTAGA 2887
Db 2093 ----- 2092
Qy 2888 CAGAAATGTAATTTCACTCTCAACATTTCTGTTAAGATTAACGTAGATTAAGATTCG 2947
Db 2093 ----- 2092
Qy 2948 CTTAGTGTGCTTGTCCAACTTTTCTTCTGATTTTCTTTTCGATTTTGAAGGTCA 3007
Db 2093 -----TAGGTCA 2100
Qy 3008 GTCAGAGCTGACGATTCAGAGCACTTCCAGAAATGATGTAGAGTGCAGAGAAATATA 3067
Db 2101 GTCAGAGCTGACGATTCAGAGCACTTCCAGAAATGATGTAGAGTGCAGAGAAATATA 2160
Qy 3068 GTATCCAAAGTGGCAAGATTAATGCTGTGGCTTTTGGGCTGATCAACCGCATAGAAATG 3127
Db 2161 GTATCCAAAGTGGCAAGATTAATGCTGTGGCTTTTGGGCTGATCAACCGCATAGAAATG 2220
Qy 3128 TTACCAGAGGTGAAGGAAATAATCTACAATTCAATCAATGTGTGAAAACGTTGGAAAT 3187
Db 2221 TTACCAG----- 2227
Qy 3188 GATTATAGTGTGAGCTGCTGTTGATTCCTGTTATTAATAGTTTGTGATGGCGCAATGCT 3247
Db 2228 -----AGGTTTGTGATGGCGCAATGCT 2249
Qy 3248 GAAATTTGCACTGACAGAGAGCTGAAATCTGCGACAGTTGGTTTATGATTAATAC 3307
Db 2250 GAAATTTGCACTGACAGAGAGCTGAAATCTGCGACAGTTGGTTTATGATTAATAC 2309
Qy 3308 ACTGTGAACCTATCTGTGACAGTGTGACAGTCTGACAGATGGAACCCGTGCTGTGT 3367
Db 2310 ACTGTGAACCTATCTGTGACAGTGTGACAGTCTGACAGATGGAACCCGTGCTGTGT 2369
Qy 3368 GGAAGCAACTCTGAGAGAGTCTGCTTCTATCTGATTTGGTTGATCCAGAAAACAAATGC 3427
Db 2370 GGAAGCAACTCTGAGAGAGTCTGCTTCTATCTGATTTGGTTGATCCAGAAAACAAATGC 2429
Qy 3428 TACTGATGTCAAACTTACCAACAAGATTAAGAAATTTCTGTGTCAAAGTCAAGGTGAAA 3487
Db 2430 TACTGATGTCAAACTTACCAACAAGATTAAGAAATTTCTGTGTCAAAGTCAAGGTGAAA 2489
Qy 3488 AATCATGAAGGCTCTGTGTCATATATATATACATATATGATGATGATGATGATGATG 3547
Db 2490 AATCATGAAGGCTCTGTGTCATATATATATATACATATATGATGATGATGATGATG 2549
Qy 3548 CGAAGATCTCTTGTCTGTAAATCTCTCTAATGTTAGTGTATTAATTAATGAACACAAA 3607
Db 2550 CGAAGATCTCTTGTCTGTAAATCTCTCTAATGTTAGTGTATTAATTAATGAACACAAA 2609
Qy 3608 AAAATTAACGTTCTTGCAACACCTTTTCTGTGATCTTAAACTATATCAATTAAGGCTTCAA 3667
Db 2610 AAAATTAACGTTCTTGCAACACCTTTTCTGTGATCTTAAACTATATCAATTAAGGCTTCAA 2669

RESULT: 4
ADJ38210 standard; DNA; 2637 BP.
ID ADJ38210

XX ADJ38210;
AC
XX 06-MAY-2004 (first entry)
DT
XX Plastid division-related Arc6 orthologue gene 3.
DE
XX prokaryotic type; plastid division; Ftn2; ARC6; ARC5; Fzo; plant cell;
KM agronomic; horticultural; crop plant; ornamental plant; woody plant;
KW herbicide target; gene; ds.
OS
XX Arabidopsis thaliana.
XX
XX WO2004001003-A2.
XX
XX 31-DEC-2003.
XX
XX 20-JUN-2003; 2003WO-US019536.
XX
XX 20-JUN-2002; 2002US-0390140P.
XX PR 09-AUG-2002; 2002US-0402242P.
XX PR 20-JUN-2003; 2003US-00600070.
XX
XX (UNMS) UNIV MICHIGAN STATE.
XX
XX Osheryoung KW, Vittha S, Koksharova OA, Gao H;
PI
XX WPI: 2004-082486/08.
XX P-PSDB; ADJ38209.
XX
XX New isolated Ftn2, ARC5 and/or Fzo-like nucleic acid sequences, useful
PT for further characterizing plastid division in plant cells, and in
PT varying agronomic and horticultural characteristics of economically
PT important plants.
XX
XX Disclosure; Fig 8; 287pp; English.
XX
XX This invention relates to novel prokaryotic type or plastid division and
CC related genes and proteins. In particular, the invention relates to novel
CC Ftn2 (ARC6), ARC5 and Fzo-like genes and polypeptides. The methods and
CC compositions of the present invention are useful for further
CC characterizing plastid division in plant cells, in order to vary
CC agronomic and horticultural characteristics of economically important
CC plants, such as crop, ornamental and woody plants. They can also be used
CC as herbicide targets. The present sequence is that of a gene which is
CC related to the invention.
XX
XX Sequence 2637 BP; 706 A; 535 C; 644 G; 752 T; 0 U; 0 Other;
SQ
Query Match 52.7%; Score 1931; DB 12; Length 2637;
Best Local Similarity 80.4%; Pred. No. 0;
Matches 2616; Conservative 0; Mismatches 5; Indels 632; Gaps 5;

QY 668 CCGCCACCAACCACCGCACTCTGCTCTCTGCCACCATCTATTGATGTCGCCAAGCC 727
DB 301 CCGCCACCAACCACCGCACTCTGCTCTGCCACCATCTATTGATGTCGCCAAGCC 360
QY 728 ACGTCCCATCCCATGATTTCTACAGAGATATTAGAGGCTCAAGCACTTTCTTAACG 787
DB 361 ACGTCCCATCCCATGATTTCTACAGAGATATTAGAGGCTCAAGCACTTTCTTAACG 420
QY 788 ATGGAATCAGAAAGAGATTCGAAAGCTAGAGGTTTCGAACCGCGCAATTCGGTTACGC 847
DB 421 ATGGAATCAGAAAGAGATTCGAAAGCTAGAGGTTTCGAACCGCGCAATTCGGTTACGC 480
QY 848 ACGACGCTTTATATACCGCGAGACAGATTTCTCAAGCTGCTTGGAATCTGTCTATC 907
DB 481 ACGACGCTTTATATACCGCGAGACAGATTTCTCAAGCTGCTTGGAATCTGTCTATC 540
QY 908 CTGGGCTAGAAAGATAGCAATGAAGGTCCTCTGTAAGAGAGAGCTACAGTCATCA 967
DB 541 CTGGGCTAGAAAGATAGCAATGAAGGTCCTCTGTAAGAGAGAGCTACAGTCATCA 600
QY 968 CTGATGTTCCCTTGAGATAGGTAATTCGATTCGAAATATAAGTTCTTCGTTTAA 1027
DB 601 CTGATGTTCCCTTGAGATAGGTAATTCGATTCGAAATATAAGTTCTTCGTTTAA 617
QY 1028 TTTCATGAATTCGATTAAGAGAGAACTTTATCTAGTGAAGTTCTCGGCGCTCTGT 1087
DB 618 TTTCATGAATTCGATTAAGAGAGAACTTTATCTAGTGAAGTTCTCGGCGCTCTGT 638
QY 1088 GTATTGCAAGAGGTGGTGAAGTGAAGTATGTTCTTGCGGTGGTGAAGGCTGCTTAAG 1147
DB 639 GTATTGCAAGAGGTGGTGAAGTGAAGTATGTTCTTGCGGTGGTGAAGGCTGCTTAAG 698
QY 1148 GAGAGGTTGCTTAAGTGTGTTAAGCAAGATGTTGTTAGTTATGCGCTTGCTTCTC 1207
DB 699 GAGAGGTTGCTTAAGTGTGTTAAGCAAGATGTTGTTAGTTATGCGCTTGCTTCTC 758
QY 1208 GATGTCGAGGAGTGTGTTAGGCAATTCGATTCACCTGATTTTATCTGTTAGATT 1267
DB 759 GATGTCGAGGAGTGTGTTAGGCAATTCGATTCACCTGATTTTATCTGTTAGATT 818
QY 1268 GTTGAGGAGCTTTGAAGCTTTTACAGGTAGTTGACTTCTTGTTGTAATTTGACGAGC 1327
DB 819 GTTGAGGAGCTTTGAAGCTTTTACAGGTAGTTGACTTCTTGTTGTAATTTGACGAGC 843
QY 1328 TTGGCTTTATAGAATCTTCTGATTTGATATCTTGTATTTGAGTCTGTGTAAGAGAA 1387
DB 844 TTGGCTTTATAGAATCTTCTGATTTGATATCTTGTATTTGAGTCTGTGTAAGAGAA 851
QY 1388 GAGCAAGTAGCCTTGACCGGATTTACGTGCACAAATGATGAGACTTTGGAAGATC 1447
DB 852 GAGCAAGTAGCCTTGACCGGATTTACGTGCACAAATGATGAGACTTTGGAAGATC 911
QY 1448 ACTCCGGTATGCTTGAAGTACTTGAGCTTCCGCTTGCTGATGATTAAGCTGCGAAA 1507
DB 912 ACTCCGGTATGCTTGAAGTACTTGAGCTTCCGCTTGCTGATGATTAAGCTGCGAAA 971
QY 1508 AGACTAATGTTTAAAGCGGTGCGGAATATTTTGTGCTGTGTGAGAGAGGTGAGCA 1567
DB 972 AGACTAATGTTTAAAGCGGTGCGGAATATTTTGTGCTGTGTGAGAGAGGTGAGCA 1031
QY 1568 TCAGCTTTGTGGGGGTTGACCGGTGAGAGATTTATGATGAGGCTTTTACGAATG 1627
DB 1032 TCAGCTTTGTGGGGGTTGACCGGTGAGAGATTTATGATGAGGCTTTTACGAATG 1091
QY 1628 ACAAGCTGAGAGAGATACAGTTTAACTTTTAAATTTCTTAGCATGATATA 1687
DB 1092 ACAAGCTGAGAGAGATACAGTTTAACTTTTAAATTTCTTAGCATGATATA 1104
QY 1688 ACTTAGTTCATTTTATGATGATGTTGTTGTTAGTTGATCTTTTGTAGCTACCC 1747
DB 1105 ACTTAGTTCATTTTATGATGATGTTGTTGTTAGTTGATCTTTTGTAGCTACCC 1128
QY 1748 CAAGCAATATTCCAGCAGAGTCAATTGAAGTTTACGAAGTTGCACTTGCTTGTCGCTC 1807

Db	1129	CAAGCAATATTCGACGACAGTCATTGGAAGTTTACGAAGTTGCACTTCTCTGTGGCTC	1168
Qy	1808	AAGCTTTATTTGGTTAAGAAAGCCACACCTTTTACAGAGATGCTGATPAGCAATTCACGCAAC	1867
Db	1189	AAGCTTTATTTGGTTAAGAAAGCCACACCTTTTACAGAGATGCTGATPAGCAATTCACGCAAC	1248
Qy	1868	TTGACGAGGCTAAGGTAATGGCTAATGGAGATTTCCGAGATGGTAATGATPACAGGAATA	1927
Db	1249	TTGACGAGGCTAAGGTAATGGCTAATGGAGATTTCCGAGATGGTAATGATPACAGGAATA	1308
Qy	1928	ATTGGAGATPAGACTTCCGTCAGAAAGGGAATCTGTGCACTGCTTATPAGCAAAAGTTG	1987
Db	1309	ATTGGAGATPAGACTTCCGTCAGAAAGGGAATCTGTGCACTGCTTATPAGCAAAAGTTG	1368
Qy	1988	ATGAAATGCCATATGTGTGTGGGCTTACGACATGAGAGATTCACATATPAGAAATCCAGCTA	2047
Db	1369	ATGAAATGCCATATGTGTGTGGGCTTACGACATGAGAGATTCACATATPAGAAATCCAGCTA	1428
Qy	2048	TTGATGAGATTTGTTTGGAGAATTCMAATTCGATGACAAATGATGATCCCTGGACAT	2107
Db	1429	TTGATGAGATTTGTTTGGAGAATTCMAATTCGATGACAAATGATGATCCCTGGACAT	1488
Qy	2108	GCAAAATTTGTGAAAACCTGTGTGGCAGGGGTGTCTTTCCTAAGTTCCAGACAACAAG	2167
Db	1489	GCAAAATTTGTGAAAACCTGTGTGGCAGGGGTGTCTTTCCTAAGTTCCAGACAACAAG	1548
Qy	2168	ATTAATAAATTTTAAACTCGGGGACATCTATGATGATCCATAGGTTTGAATCTTTGAAA	2227
Db	1549	ATTAATAAATTTTAAACTCGGGGACATCTATGATGATCCATAGGTTTGAATCTTTGAAA	1608
Qy	2228	GAGTGGAGATGTTTCAAGGGTTCTCCCTTAGCGCTGCTGCAACTATGCGCAAGAAATGGAG	2287
Db	1609	GAGTGGAGATGTTTCAAGGGTTCTCCCTTAGCGCTGCTGCAACTATGCGCAAGAAATGGAG	1668
Qy	2288	CCGAGCATGTGAAAGCTATAGTGCCTATGACAGGACATGCAAGAAATTTTCTTCCCGCTATA	2347
Db	1669	CCGAGCATGTGAAAGCTATAGTGCCTATGACAGGACATGCAAGAAATTTTCTTCCCGCTATA	1728
Qy	2348	CAGATPAGAAACTCGGCTGACACCCCAAGATGTGTGCAAGAGACAGTGTTTAGTATGATCCTG	2407
Db	1729	CAGATPAGAAACTCGGCTGACACCCCAAGATGTGTGCAAGAGACAGTGTTTAGTATGATCCTG	1788
Qy	2408	TTGGTTAACAAATGTAGGCGCGTATGAGTGTGAGCCTGTGTTTATTTGCAAGAACTGTAAAGAC	2467
Db	1789	TTGGTTAACAAATGTAGGCGCGTATGAGTGTGAGCCTGTGTTTATTTGCAAGAACTGTAAAGAC	1848
Qy	2468	CCTCTGAAAACCTTTGAAACTATATGATTTATGCAATTTGAGCTGGGGTCTCAGAGATGACG	2527
Db	1849	CCTCTGAAAACCTTTGAAACTATATGATTTATGCAATTTGAGCTGGGGTCTCAGAGATGACG	1908
Qy	2528	TTGATPAGAAACTACGTGTGAAATGTCCGTGTGATATGTTTAAAGAGGCAAGTGTGAAGA	2587
Db	1909	TTGATPAGAAACTACGTGTGAAATGTCCGTGTGATATGTTTAAAGAGGCAAGTGTGAAGA	1968
Qy	2588	TTCTTAGCTGCTGCTGTGTGGCAATTTGGAATCTGATTTCACTGTGACGCCAGAGATTTTCTTA	2647
Db	1969	TTCTTAGCTGCTGCTGTGTGGCAATTTGGAATCTGATTTCACTGTGACGCCAGAGATTTTCTTA	2028
Qy	2648	AAAGCAAGCTACATTTTTCACGCAAGAGATATGTTTTCTTCTATGGAATCTGATGTGCTA	2707
Db	2029	AAAGCAAGCTACATTTTTCACGCAAGAGATATGTTTTCTTCTATGGAATCTGATGTGCTA	2088
Qy	2708	CCATATGATATGATTTAAATGATGCAATTTTCAATATCTGATGTGCTCAAAATATGCTTGT	2767
Db	2089	CCA-----	2091
Qy	2768	TTGTGTGAGCTAAGAACATATGTTCCACTTATATACATGTCACAAAAGTTGTACCAAGATTA	2827
Db	2092	-----	2091
Qy	2828	ACAAGTGTGCTGAGTAAATTTTCACTAATATATGCTGCTGAAATTTTGTATCAAACTGTAGA	2887

Db	2092	-----	2091
QY	2888	CAGAAATGTAATTCTCACTGCAACATTCTGTTAGAAATACGTAGGATTAGAGATTGC	2947
Db	2092	-----	2091
QY	2948	CTTAGTGTGGCTTTGTCCAACTTTCTTCCTTGATTCTTTTTCGATTTAGGGTCA	3007
Db	2092	-----	2099
QY	3008	GTCCAGACTGCAGATTGCAGAGACCTTCCAGATGAGTCTAGACGTGCAGAGAAATATA	3067
Db	2100	GTCCAGACTGCAGATTGCAGAGACCTTCCAGATGAGTCTAGACGTGCAGAGAAATATA	2159
QY	3068	GTATCCAAAGTGGCAGAGAAATTAAGTCTCTGGCTTTTGGGCTTATCACCCGATAGAAATG	3127
Db	2160	GTATCCAAAGTGGCAGAGAAATTAAGTCTCTGGCTTTTGGGCTTATCACCCGATAGAAATG	2219
QY	3128	TTACCAAGGTGAGGGAATTAATCTCAATTCATTCATTTGTGTGAAAACTGTTGGACAT	3187
Db	2220	TTACCAAG	2226
QY	3188	GATTATAGTCTGGTCCCTGTTTGATTCTGTATTTATAGTTTGGATGGCGAATGCT	3247
Db	2227	-----	2248
QY	3248	GAAGATTTTGGACTGCAGACAGCAGCTGAACTGCGACCTTGGGTTGGTTTATGATTATAC	3307
Db	2249	GAAGATTTTGGACTGCAGACAGCAGCTGAACTGCGACCTTGGGTTGGTTTATGATTATAC	2308
QY	3308	ACTGTTGAAAATCTATCTGTTTGAACAGTGTGACAGTCTCAGCGAGATGGAACCCGTGCTCTGGT	3367
Db	2309	ACTGTTGAAAATCTATCTGTTTGAACAGTGTGACAGTCTCAGCGAGATGGAACCCGTGCTCTGGT	2368
QY	3368	GGAAGCAACTCTGGAGGAGTCTGCTGTTCTATCTGATTTGGTTCATCCAGAAAACAATGC	3427
Db	2369	GGAAGCAACTCTGGAGGAGTCTGCTGTTCTATCTGATTTGGTTCATCCAGAAAACAATGC	2428
QY	3428	TACTGATGTGAGAACTTACCAACAAGATACGAATTTTCTGGTCCAAAGTCAGGGTGGAA	3487
Db	2429	TACTGATGTGAGAACTTACCAACAAGATACGAATTTTCTGGTCCAAAGTCAGGGTGGAA	2488
QY	3488	AATACACTGAAAGGCTCTGTTCTTGGCATATATATATCTCATATGTAGCATGTCTGAGCTTG	3547
Db	2489	AATACACTGAAAGGCTCTGTTCTTGGCATATATATATCTCATATGTAGCATGTCTGAGCTTG	2548
QY	3548	CGAGATCTCTTTGTTCTGTTAAATCTCTCAATGCTAGTGTGTTAAATGAACAACAA	3607
Db	2549	CGAGATCTCTTTGTTGTTTAAATCTCTCTCAATGTTAGTGTGTTAAATGAACAACAA	2608
QY	3608	AAAATTACGTTTC	3620
Db	2609	AAAATTACGTTTC	2621
RESULT 5			
ADJ38129 standard; cDNA; 2406 BP.			
ADJ38129			
ADJ38129;			
06-MAY-2004 (first entry)			
Arabidopsis thaliana Arc6-1 cDNA SeqID1.			
prokaryotic type; plaetid division; Ftn2; ARC6; ARCF; Fzo; plant cell;			
agronomic; horticultural; crop plant; ornamental plant; woody plant;			
herbicide target; gene; ss.			
Arabidopsis thaliana.			
WO2004001003-A2.			

PD 31-DEC-2003.
XX
XX 20-JUN-2003; 2003WO-US019536.
XX
PR 20-JUN-2002; 2002US-0390140P.
PR 09-AUG-2002; 2002US-0402242P.
PR 20-JUN-2003; 2003US-00600070.
XX
PA (UNMS) UNITV MICHIGAN STATE.
XX
PI Oseeryoung KW, Vittha S, Kokeharova OA, Gao H;
XX
DR WPI; 2004-082486/08.
DR P-PSDB; ADJ38202.
XX
XX New isolated Flt2, ARCS and/or Pzo-like nucleic acid sequences, useful
PT for further characterizing plastid division in plant cells, and in
PT varying agronomic and horticultural characteristics of economically
PT important plants.
XX
PS Claim 1; SEQ ID NO 1; 287bp; English.
XX
CC This invention relates to novel prokaryotic type or plastid division and
CC related genes and proteins, in particular, the invention relates to novel
CC Flt2 (ARCS), ARCS and Pzo-like genes and polypeptides. The methods and
CC compositions of the present invention are useful for further
CC characterizing plastid division in plant cells, in order to vary
CC agronomic and horticultural characteristics of economically important
CC plants, such as crop, ornamental and woody plants. They can also be used
CC as herbicide targets. The present sequence is a cDNA sequence which is
CC related to the invention.
XX
SQ Sequence 2406 BP; 612 A; 492 C; 618 G; 684 T; 0 U; 0 Other;
Query March 47.0%; Score 1724; DB 12; Length 2406;
Best Local Similarity 79.2%; Pred. No. 0;
Matches 2406; Conservative 0; Mismatches 0; Indels 632; Gaps 5;
QY 481 ATGGAAGCTGAGTCAAGTCCGACATTTGCTCTCCCATTCATATATGCGGATTACCA 540
DB 1 ATGGAAGCTGAGTCAAGTCCGACATTTGCTCTCCCATTCATATATGCGGATTACCA 60
QY 541 CCGGCGACGACAAAGCTCCGACGTAGCCAAACACCTTCAACATATTTGCTCCGACGC 600
DB 61 CCGGCGACGACAAAGCTCCGACGTAGCCAAACACCTTCAACATATTTGCTCCGACGC 120
QY 601 AAAAGGCGGACCGCTCTCTCCGACCTTCAATTTGACCTCCGATTCCTCTCTCTCC 660
DB 121 AAAAGGCGGACCGCTCTCTCCGACCTTCAATTTGACCTCCGATTCCTCTCTCTCC 180
QY 661 TTGCGCACCGCACCAACCGCCACTCTGCTCTCTGCGACCATTAATGATGCTGCC 720
DB 181 TTGCGCACCGCACCAACCGCCACTCTGCTCTCTGCGACCATTAATGATGCTGCC 240
QY 721 GAAAGCGACGTCCTCCCATTCGATTTCTACCGAGTATTAGAGCTCAAAACATTTG 780
DB 241 GAAAGCGACGTCCTCCCATTCGATTTCTACCGAGTATTAGAGCTCAAAACATTTG 300
QY 781 TTAACCGATGGAATCAGAAAGACATTTGAGGTAGGTTTGAAGAACCGCGCAATTCGCT 840
DB 301 TTAACCGATGGAATCAGAAAGACATTTGAGGTAGGTTTGAAGAACCGCGCAATTCGCT 360
QY 841 TTGACGACGACGCTTTAATCAGCGGACGACAGATTTCTCAAGCTGCTTGCGAAACTGTG 900
DB 361 TTGACGACGACGCTTTAATCAGCGGACGACAGATTTCTCAAGCTGCTTGCGAAACTGTG 420
QY 901 TCTAATCTCGGTCTTGAAGAGAGTACATGAAAGTCTTTTGATGATGAAGAACTACA 960
DB 421 TCTAATCTCGGTCTTGAAGAGAGTACATGAAAGTCTTTTGATGATGAAGAACTACA 480
QY 961 GTCAATCACTGATGTTCTTTGGGATPAAGGTAATTCGATTTGGGAATATAAGTTTCTTC 1020
DB 481 GTCAATCACTGATGTTCTTTGGGAT----- 504

QY 1021 GTTTTAATTTCAATGATTTGATTAAGGAGAACTTTATCTAGTGAAGTTCCTGGGCG 1080
DB 505 -----AAGGTTCTGGGCG 518
QY 1081 TCTCTGTGTATTGCAAGAAAGTGTGACTGATGATGATGTTCTTCGGGTTGGTAGGCTCT 1140
DB 519 TCTCTGTGTATTGCAAGAAAGTGTGACTGATGATGATGTTCTTCGGGTTGGTAGGCTCT 578
QY 1141 GCTTAAGAGAGGTTGCTTAAGTCTTAAAGCAAGATGCTTTTAAAGTGAAGGCGCTTG 1200
DB 579 GCTTAAGAGAGGTTGCTTAAGTCTTAAAGCAAGATGCTTTTAAAGTGAAGGCGCTTG 638
QY 1201 GTTTCTGATGTCTGAGGAGATGCTATGCGATTGATCCACTGATTTTAATTAACGTGTTA 1260
DB 639 GTTTCTGATGTCTGAGGAGATGCTATGCGATTGATCCACTGATTTTAATTAACGTGTTA 698
QY 1261 TGAATTTGTTGAGAGACTTTGAAGCTTTTACAGGTTAGTTGACTGCTTTGGTAATTTG 1320
DB 699 TGAATTTGTTGAGAGACTTTGAAGCTTTTAC----- 730
QY 1321 ACAGAGCTTGGCTTTAAGAACTTTCTGATTTGATTAATTTGATGATCTTGCTGA 1380
DB 731 -----A 731
QY 1381 GAGAGAAAGAGCAATGATGCTTGACCGGATTTACGTGCAAAATTTGATGACTTTGA 1440
DB 732 GAGAGAAAGAGCAATGATGCTTGACCGGATTTACGTGCAAAATTTGATGACTTTGA 791
QY 1441 AGAGATCACTCCGCTTATGCTTTGAGCTACCTTGGCTTACCGCTTGGTATGATTAACGC 1500
DB 792 AGAGATCACTCCGCTTATGCTTTGAGCTACCTTGGCTTACCGCTTGGTATGATTAACGC 851
QY 1501 TGGAAAAAGACTAATGTTTAAAGCGGTGCGGAATATTTTGGTCTGTGAGAGAGG 1560
DB 852 TGGAAAAAGACTAATGTTTAAAGCGGTGCGGAATATTTTGGTCTGTGAGAGAGG 911
QY 1561 TGGAGCATCAGCTTTGTTGGGGTTTGACCCGTGAGAACTTTATGATGAGGCGTTT 1620
DB 912 TGGAGCATCAGCTTTGTTGGGGTTTGACCCGTGAGAACTTTATGATGAGGCGTTT 971
QY 1621 ACCGAATACAGCTGCTGAGAGGATTAAGATTAATCTTTTAAATTTCTTTAGCA 1680
DB 972 ACCGAATACAGCTGCTGAG----- 991
QY 1681 TGAATATACTTAAGTTTCAATTTAATGATGTTGAGTGAATCTTTTGTGA 1740
DB 992 -----AGGTTGATCTTTTGTGA 1008
QY 1741 GCTAACCCCAAGCAATATTCAGAGAGTCAATTTGAAGTTTACGAAGTTGACCTTGCTTT 1800
DB 1009 GCTAACCCCAAGCAATATTCAGAGAGTCAATTTGAAGTTTACGAAGTTGACCTTGCTTT 1068
QY 1801 GTGGCTCAAGCTTTATTTGTTGAAGAGCCACACTTTTAAGAGATGCTGATTAAGCAATTC 1860
DB 1069 GTGGCTCAAGCTTTATTTGTTGAAGAGCCACACTTTTAAGAGATGCTGATTAAGCAATTC 1128
QY 1861 CAGCAACTTTCAGAGGTTAAGTATGCTATGAGAGATTCCTGCAATGTTGATATACA 1920
DB 1129 CAGCAACTTTCAGAGGTTAAGTATGCTATGAGAGATTCCTGCAATGTTGATATACA 1188
QY 1921 CCGAATTAATTTGGAGATTAAGTTCGCTTGAAGAAAGGGAATCTGTGCACTGCTTTATAGGC 1980
DB 1189 CCGAATTAATTTGGAGATTAAGTTCGCTTGAAGAAAGGGAATCTGTGCACTGCTTTATAGGC 1248
QY 1981 AAAGTTATGATGATGCTGATGCTTGGGCTTGAAGATGATGATTAAGCAATTAAGCAAT 2040
DB 1249 AAAGTTATGATGATGCTGATGCTTGGGCTTGAAGATGATGATTAAGCAATTAAGCAAT 1308
QY 2041 CCAAGCTATTTGAGAGTTGTTTGAAGAAATTCAAATGCTATGATCAATGATGATCTCCCT 2100
DB 1309 CCAAGCTATTTGAGAGTTGTTTGAAGAAATTCAAATGCTATGATCAATGATGATCTCCCT 1368

Qy	2101	GGACATATCAAAATTTGTTGAAACCGGTGTGGACGGGTGTCTTCCAGGTTCAAGAC	2160
Db	1369	GGACATATCAAAATTTGTTGAAACCGGTGTGGACGGGTGTCTTCCAGGTTCAAGAC	1428
Qy	2161	ACCAAAAGTAAAAATTTTAACTCGGGGACTACTATGATGATTCCTATGTTTGAATTAC	2220
Db	1429	ACCAAAAGTAAAAATTTTAACTCGGGGACTACTATGATGATTCCTATGTTTGAATTAC	1488
Qy	2221	TTGGAAAAGTGGAGAGTATGTTCAAGGTTCTCTCTTAACTGCTGTCGCACTATGGCAAG	2280
Db	1489	TTGGAAAAGTGGAGAGTATGTTCAAGGTTCTCTCTTAACTGCTGTCGCAACTATGGCAAG	1548
Qy	2281	ATTGGAGCCGAGCAGTGTAAAGCTATAGTCTATGACGAGCAGTGCAGAAAGTTTTCTTCC	2340
Db	1549	ATTGGAGCCGAGCAGTGTAAAGCTATAGTCTATGACGAGCAGTGCAGAAAGTTTTCTTCC	1608
Qy	2341	CGCTATACAGATAGTAAACTCGGGCTGAACCCAGAGATGTGCAGAGACAGTGTTAGTATA	2400
Db	1609	CGCTATACAGATAGTAAACTCGGGCTGAACCCAGAGATGTGCAGAGACAGTGTTAGTATA	1668
Qy	2401	GATCCCTGTGGTAAACAATGTATAGGCCGTGTAGTGTGACCTGTGTCTTTATTTGCAGAACT	2460
Db	1669	GATCCCTGTGGTAAACAATGTATAGGCCGTGTAGTGTGACCTGTGTCTTTATTTGCAGAACT	1728
Qy	2461	GTAAGACCCCTGTGAAAACCTTTGAAAACCTATATATATGCAATTTCGAGCTGGGGGTCTCAGAG	2520
Db	1729	GTAAGACCCCTGTGAAAACCTTTGAAAACCTATATATATGCAATTTCGAGCTGGGGGTCTCAGAG	1788
Qy	2521	AGTAGCCGTGTATGAAACTACTAGTTGAAATGTCCGTCGTGATGTATGTTAAAGAGGCAAGT	2580
Db	1789	AGTAGCCGTGTATGAAACTACTAGTTGAAATGTCCGTCGTGATGTATGTTAAAGAGGCAAGT	1848
Qy	2581	GTGAAGATCCTAGCTGTGCTGTGTGTGGCAATTGAGCTGATTTTCACTGTTCAGCCAGAAATAT	2640
Db	1849	GTGAAGATCCTAGCTGTGCTGTGTGTGGCAATTGAGCTGATTTTCACTGTTCAGCCAGAAATAT	1908
Qy	2641	TTTTCTTAAAGACAGCTCATCTTTTTCACAGCAGATATGTGTTTCTTTATATAGGATCTGAT	2700
Db	1909	TTTTCTTAAAGACAGCTCATCTTTTTCACAGCAGATATGTGTTTCTTTATATAGGATCTGAT	1968
Qy	2701	GTGCGTACACATAGGATGATTTAAATGATGCAATTTTCAATATATCTGCATTTGCTCAAAATA	2760
Db	1969	GTGCGTACACACATAGGATGATTTAAATGATGCAATTTTCAATATATCTGCATTTGCTCAAAATA	1978
Qy	2761	TGCTTGTGTTGTGAGCTAAGAAACATATGTTCCACTAATATACATGTCCCAAAAGTTGTACC	2820
Db	1979	TGCTTGTGTTGTGAGCTAAGAAACATATGTTCCACTAATATACATGTCCCAAAAGTTGTACC	1978
Qy	2821	AAGATTTAACAAAGTTGCTGAGTAAATTTTCACTAATATATAGCTGCTGAATTTTTTGATCAAA	2880
Db	1979	AAGATTTAACAAAGTTGCTGAGTAAATTTTCACTAATATATAGCTGCTGAATTTTTTGATCAAA	1978
Qy	2881	CTGTAGACAGAAATGTATAATTTCACTCTCAACATTTCTGTTTAAGATTAAGCTAGATTAAG	2940
Db	1979	CTGTAGACAGAAATGTATAATTTCACTCTCAACATTTCTGTTTAAGATTAAGCTAGATTAAG	1978
Qy	2941	AGATTGCTTATGTGTGCTGCTTGTGCCAACTTTCTTCCCTGATTTTCTTTTCGATTT	3000
Db	1979	AGATTGCTTATGTGTGCTGCTTGTGCCAACTTTCTTCCCTGATTTTCTTTTCGATTT	1979
Qy	3001	AGGCTCAGTCAGAGCTGACGATTCAGAGACCTTCCAGAAATGAGTGTAGAGCTGACAG	3060
Db	1980	AGGCTCAGTCAGAGCTGACGATTCAGAGACCTTCCAGAAATGAGTGTAGAGCTGACAG	2039
Qy	3061	GAATATATGATATCCAAAGTGGCAGAAAGATTAAATCTCTGCTTTTGGGCTGATACCCGAT	3120
Db	2040	GAATATATGATATCCAAAGTGGCAGAAAGATTAAATCTCTGCTTTTGGGCTGATACCCGAT	2099
Qy	3121	AGAAATGTTACAGAGGTGAGGGAGATAAATCTACAAATTCATCAATGATGTGGAACAACTGT	3180
Db	2100	AGAAATGTTACAG-----	2113
Qy	3181	TGGAACATGATTAATAGTCTGTGCTTGTGATTTCTGTATTTAATAGGTTTGGATGGC	3240

Db	2114	-----		AGTTTGGATGGC	2128
Qy	3341	GAATGCTGAAGATTGGACTGACAGACAGCTGAACTGCGCAGCTTGAGTTGTTATG			3300
Db	2129	GAATGCTGAAGATTGGACTGACAGACAGCTGAACTGCGCAGCTTGAGTTGTTATG			2188
Qy	3301	ATTATACACGTGTTAACTATCTGTTTGAACGTGTGACAGTCTCAGCAGATGMAACCCGTG			3360
Db	2189	ATTATACACGTGTTAACTATCTGTTTGAACGTGTGACAGTCTCAGCAGATGMAACCCGTG			2248
Qy	3361	CTCTGGTGAAGCAACTCTGGAAGAGCTGCTGTCATCTGATTTGGTTCATCCAGAAA			3420
Db	2249	CTCTGGTGAAGCAACTCTGGAAGAGCTGCTGTCATCTGATTTGGTTCATCCAGAAA			2308
Qy	3421	ACAACTGCTACTGATGTGACAACTCACAACAAGATACGAAGTTTCTGCTCAAGTCAG			3480
Db	2309	ACAACTGCTACTGATGTGACAACTCACAACAAGATACGAAGTTTCTGCTCAAGTCAG			2368
Qy	3481	GGTGGAAAATCCTCAGGAGCTGTGTTCTTGATCATATTA			3518
Db	2369	GGTGGAAAATCCTCAGGAGCTGTGTTCTTGATCATATTA			2406
RESULT 6					
ADJ38135	ADJ38135	standard; cDNA; 2406 BP.			
XX	ADJ38135;				
XX	06-MAY-2004	(first entry)			
XX	Arabidopsis thaliana AtFtn2 cDNA SegID9.				
XX	prokaryotic type; plastid division; Ftn2; ARC6; ARC5; Fzo; plant cell;				
XX	agronomic; horticultural; crop plant; ornamental plant; woody plant;				
XX	herbicide target; gene; ss.				
XX	Arabidopsis thaliana.				
XX	MO2004001003-A2.				
XX	31-DEC-2003.				
XX	20-JUN-2003; 2003MO-US019536..				
XX	20-JUN-2002; 2002US-0390140P.				
XX	09-AUG-2002; 2002US-0402242P.				
XX	20-JUN-2003; 2003US-0060070.				
XX	(UNMS) UNIV MICHIGAN STATE.				
XX	OserYoung KW, Vicha S, Koksharova OA, Gao H;				
XX	WI; 2004-082486/08.				
XX	P-PSDB; ADJ38203.				
XX	New isolated Ftn2, ARC5 and/or Fzo-like nucleic acid sequences, useful				
XX	for further characterizing plastid division in plant cells, and in				
XX	varying agronomic and horticultural characteristics of economically				
XX	important plants.				
XX	Claim 1; SEQ ID NO 9; 287bp; English.				
XX	This invention relates to novel prokaryotic type or plastid division and				
XX	related genes and proteins. In particular, the invention relates to novel				
XX	Ftn2 (ARC6), ARC5 and Fzo-like genes and polypeptides. The methods and				
XX	compositions of the present invention are useful for further				
XX	characterizing plastid division in plant cells, in order to vary				
XX	agronomic and horticultural characteristics of economically important				
XX	plants, such as crop, ornamental and woody plants. They can also be used				
XX	as herbicide targets. The present sequence is a cDNA sequence which is				
XX	related to the invention.				

Qy 841 TTCACGACGACGCTTTAATCATGCCGAGACAGATTCTTCAAGCTGCTTGGCAAACCTGCG 900
Db 361 TTCACGACGACGCTTTAATCATGCCGAGACAGATTCTTCAAGCTGCTTGGCAAACCTGCG 420
Qy 901 TCTAATCTCGGCTCTAAGAAAGATACATGAAGCTCTTCTTGATGATGAAGAACTACA 960
Db 421 TCTAATCTCGGCTCTAAGAAAGATACATGAAGCTCTTCTTGATGATGAAGAACTACA 480
Qy 961 GTCATCACTGATGCTCTTGGGATAAGGTAATTCGATTCGGAAATTAAGATTCCTTC 1020
Db 481 GTCATCACTGATGCTCTTGGGAT----- 504
Qy 1021 GTTTAAATTCATGAATGGATTAAGAAAGAACTTTATCTAGTGAAGCTTCTCGGAC 1080
Db 505 -----AAGTTCCTGCTGCG 518
Qy 1081 TCTCTGTATTTGCAAGAAAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1140
Db 519 TCTCTGTATTTGCAAGAAAGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 578
Qy 1141 GCTTAAGAGAGAGTTCCTAAGTCTTTAAGCAAGATGATGATGATGATGATGATGATGATGAT 1200
Db 579 GCTTAAGAGAGAGTTCCTAAGTCTTTAAGCAAGATGATGATGATGATGATGATGATGATGAT 638
Qy 1201 GTTTCGATGCTCTGAGGAGTGTATGACATGATGATGATGATGATGATGATGATGATGAT 1260
Db 639 GTTTCGATGCTCTGAGGAGTGTATGACATGATGATGATGATGATGATGATGATGATGAT 698
Qy 1261 TGAAGTTTGTAGAGAGCTTGAAGCTTTACAGATGATGATGATGATGATGATGATGATGAT 1320
Db 699 TGAAGTTTGTAGAGAGCTTGAAGCTTTAC----- 730
Qy 1321 ACGAGCTGTGGCTTATTAAGAACTTTCTGATTTGATGATGATGATGATGATGATGATGAT 1380
Db 731 -----A 731
Qy 1381 GGAGAGAGAGCAAGTATGACCTTGTGACCCGATTTTACGTGACAAATGTATGATGATGATGAT 1440
Db 732 GGAGAGAGAGCAAGTATGACCTTGTGACCCGATTTTACGTGACAAATGTATGATGATGATGAT 791
Qy 1441 AGAGATCACTCCGCTTATGCTTGTGAGAGTACTTGGCTTACCGCTTGGTGTATGATGATGAT 1500
Db 792 AGAGATCACTCCGCTTATGCTTGTGAGAGTACTTGGCTTACCGCTTGGTGTATGATGATGAT 851
Qy 1501 TGCGAAAAGACTAAATGCTTAAAGCGGTGTGCGAAATATTTTGTGTGCTGTGAGAGAG 1560
Db 852 TGCGAAAAGACTAAATGCTTAAAGCGGTGTGCGAAATATTTTGTGTGCTGTGAGAGAG 911
Qy 1561 TGAGCATCACTGCTTGTGGGGTTTGAACCGTGAGAAAGTTTATGAATGAGGCGTTTT 1620
Db 912 TGAGCATCACTGCTTGTGGGGTTTGAACCGTGAGAAAGTTTATGAATGAGGCGTTTT 971
Qy 1621 ACGAATGACAGTGTGAGCAAGTATACAGTTTATACCTTTTATTTCTTTAGCA 1680
Db 972 ACGAATGACAGTGTGAGCA----- 991
Qy 1681 TGAATTAATCTTAGGTTCTCATTTTAATGATGTTGTGTGTAGGTTGATCTTTTGTGA 1740
Db 992 -----AGTTGATCTTTTGTGA 1008
Qy 1741 GCTACCCCAAGCAATATTCACAGACAGTCAATTTGAAGTTTGAAGAGTTGACCTGCTCT 1800
Db 1009 GCTACCCCAAGCAATATTCACAGACAGTCAATTTGAAGTTTGAAGAGTTGACCTGCTCT 1068
Qy 1801 GTGGCTCAAGCTTTTATGTTAAGAAAGCAACCTTTTACAGAGATGCTGATTAAGCAATTC 1860
Db 1069 GTGGCTCAAGCTTTTATGTTAAGAAAGCAACCTTTTACAGAGATGCTGATTAAGCAATTC 1128
Qy 1861 CAGCAACTTCACAGAGCTTAAGGTATGAGATTCCTGCGAGTGTGATGATGATGATGATGAT 1920
Db 1129 CAGCAACTTCACAGAGCTTAAGGTATGAGATTCCTGCGAGTGTGATGATGATGATGATGAT 1188
Qy 1921 CGGAATTAATTTGGAGATGATGATTCGGTCTAAGAAAGGAGACTCTGTGCACTGCTTATGAGC 1980

Db 1189 CGGAATTAATTTGGAGATGATGATTCGGTCTAAGAAAGGAGACTCTGTGCACTGCTTATGAGC 1248
Qy 1981 AAAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2040
Db 1249 AAAGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1308
Qy 2041 CCAGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2100
Db 1309 CCAGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1368
Qy 2101 GGACTATGCAAAATTTGTGAAACCTGCTGTGCGAGGGGTGTCTTCTTCTAGTTTCAAGAC 2160
Db 1369 GGACTATGCAAAATTTGTGAAACCTGCTGTGCGAGGGGTGTCTTCTTCTAGTTTCAAGAC 1428
Qy 2161 ACCAAAGTAAAAAATTTTAACTCGGGGACTACTATGATGATGATGATGATGATGATGATGATGAT 2220
Db 1429 ACCAAAGTAAAAAATTTTAACTCGGGGACTACTATGATGATGATGATGATGATGATGATGATGAT 1488
Qy 2221 TTGAAAAGAGTGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2280
Db 1489 TTGAAAAGAGTGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1548
Qy 2281 ATTGAGCCGAGCATGTTGAAGAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 2340
Db 1549 ATTGAGCCGAGCATGTTGAAGAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 1608
Qy 2341 CGCTATACAGTAACTCGGCTGAACCCAGAGATGTCAGAGACAGTATGATGATGATGATGAT 2400
Db 1609 CGCTATACAGTAACTCGGCTGAACCCAGAGATGTCAGAGACAGTATGATGATGATGATGAT 1668
Qy 2401 GATCCTGTGTGTAACAATGATGAGGCGTGAATGATGATGATGATGATGATGATGATGATGATGAT 2460
Db 1669 GATCCTGTGTGTAACAATGATGAGGCGTGAATGATGATGATGATGATGATGATGATGATGATGAT 1728
Qy 2461 GTTAAGACCTCTGAAAACTTTGAACCTTAATGATGATGATGATGATGATGATGATGATGATGAT 2520
Db 1729 GTTAAGACCTCTGAAAACTTTGAACCTTAATGATGATGATGATGATGATGATGATGATGATGAT 1788
Qy 2521 AGTAGCGTGAATGAATCACTGTTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 2580
Db 1789 AGTAGCGTGAATGAATCACTGTTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 1848
Qy 2581 GTGAAGATCTTACGCTGTGTGTGCAATTTGACTGATTTTCACTGTTCAAGCAGAAATAT 2640
Db 1849 GTGAAGATCTTACGCTGTGTGTGCAATTTGACTGATTTTCACTGTTCAAGCAGAAATAT 1908
Qy 2641 TTTCTTAAAGCAGCTCATCTTTTCAACGCAAGATATGTTTCTTCTATGGAATCTGAT 2700
Db 1909 TTTCTTAAAGCAGCTCATCTTTTCAACGCAAGATATGTTTCTTCTATGGAATCTGAT 1968
Qy 2701 GTGCTACGATAGTATGATTAATGATGCAATTTTCAATATCTGATTCCTCAAAATA 2760
Db 1969 GTGCTACCA----- 1978
Qy 2761 TGCTTGTGTTGTGAGCTAAGAACATAGTCCCATTAATACATGTCCCAAAAGTTGTACC 2820
Db 1979 ----- 1978
Qy 2821 AAGATTAAACAAGTTGCTGAGTAATTTCACTAATATAGCTGCTTGAATTTTGTGATCAAA 2880
Db 1979 ----- 1978
Qy 2881 CTGTAGACAGAAATGTAATTTCACTCAACATTTCTGTTTGAATPAAGTAGATTAAG 2940
Db 1979 ----- 1978
Qy 2941 AAGATTGCTTAGTGCTGTGCTTGTCCAACTTTTCTTCTGATTTTCTTTTTCGATTT 3000
Db 1979 -----T 1979
Qy 3001 AGGTCAGTCAAGACTACGATTCAGAAAGCACTTCCAGAAATGATGCTAGAGATCGACAG 3060

Db 1980 AGGTCATGACAGCTGACGATTGACAGACCTTCCCAATGATGCTAGACGTCGACA 2039
 Qy 3061 GAATATGATGATCCAGATGACAGAGATTAAGTCTGAGCTTTGGCCCTGATCACCAGCAT 3120
 Db 2040 GAATATGATGATCCAGATGACAGAGATTAAGTCTGAGCTTTGGCCCTGATCACCAGCAT 2099
 Qy 3121 AGAATGTTTACCAAGAGGTGAGGAAATTAATCTCAATCAATGTTGTGAAAACTGT 3180
 Db 2100 AGAATGTTTACCAAGAGGTGAGGAAATTAATCTCAATCAATGTTGTGAAAACTGT 2113
 Qy 3181 TGACATGATTAATGCTGCTGCTTTGTTGATTTCTTATTTATGATTTGATGATGAGC 3240
 Db 2114 -----AGGTTTGGATGGCC 2128
 Qy 3241 GAATGCTGAAGATTTGAGCTGACAGAGCAGCTGAACTGCGAGCTTTGGTTTATG 3300
 Db 2129 GAATGCTGAAGATTTGAGCTGACAGAGCAGCTGAACTGCGAGCTTTGGTTTATG 2188
 Qy 3301 ATTATACATGTTGAAACCTATCTGTTGACAGTGTGACAGTCTGACAGATGAACTG 3360
 Db 2189 ATTATACATGTTGAAACCTATCTGTTGACAGTGTGACAGTCTGACAGATGAACTG 2248
 Qy 3361 CTCTGGTGAAGACATCTGGAAGAGTCTGCTGCTATCTGATTTGTTGATCCAGAA 3420
 Db 2249 CTCTGGTGAAGACATCTGGAAGAGTCTGCTGCTATCTGATTTGTTGATCCAGAA 2308
 Qy 3421 ACAATGCTACTGATGTCAGAACCTACACAACAAGATGAGATTTGCTGCAAGTCAG 3480
 Db 2309 ACAATGCTACTGATGTCAGAACCTACACAACAAGATGAGATTTGCTGCAAGTCAG 2368
 Qy 3481 GGTTGAAATCACTGAAAGCTCTGTTTGCATCATTA 3518
 Db 2369 GGTTGAAATCACTGAAAGCTCTGTTTGCATCATTA 2406
 RESULT 8
 ADJ38212/c
 ID ADJ38212 standard; cDNA; 561 BP.
 AC ADJ38212;
 XX
 DT 06-MAY-2004 (first entry)
 DE Plastid division-related Arc6 orthologue cDNA 5.
 XX
 KW prokaryotic type; plastid division; Ftn2; ARC6; ARCS; Fzo; plant cell;
 XX agronomic; horticultural; crop plant; ornamental plant; woody plant;
 XX herbicide target; gene; ss.
 XX
 OS Arabidopsis thaliana.
 XX
 PN WO2004001003-A2.
 PD 31-DEC-2003.
 PF 20-JUN-2003; 2003WO-US019536.
 PR 20-JUN-2002; 2002US-0390140P.
 PR 09-AUG-2002; 2002US-0402242P.
 PR 20-JUN-2003; 2003US-00600070.
 XX
 PA (UNMS) UNITV MICHIGAN STATE.
 PI Oseleyoung KW, Vitha S, Koksharova OA, Gao H;
 XX WPI; 2004-082486/08.
 DR
 XX
 PT New isolated Ftn2, ARCS and/or Fzo-like nucleic acid sequences, useful
 PT for further characterizing plastid division in plant cells, and in
 PT varying agronomic and horticultural characteristics of economically
 PT important plants.
 XX
 PS Disclosure; Fig 8; 287bp; English.

XX This invention relates to novel prokaryotic type or plastid division and
 CC related genes and proteins. In particular, the invention relates to novel
 CC Ftn2 (ARC6), ARCS and Fzo-like genes and polypeptides. The methods and
 CC compositions of the present invention are useful for further
 CC characterizing plastid division in plant cells, in order to vary
 CC agronomic and horticultural characteristics of economically important
 CC plants, such as crop, ornamental and woody plants. They can also be used
 CC as herbicide targets. The present sequence is a cDNA sequence which is
 CC related to the invention.
 XX
 SQ Sequence 561 BP; 167 A; 136 C; 107 G; 148 T; 0 U; 3 Other;
 Query Match 11.0%; Score 402.2; DB 12; Length 561;
 Best Local Similarity 82.2%; Pred. No. 7e-94;
 Matches 514; Conservative 0; Mismatches 20; Indels 91; Gaps 1;
 Qy 2970 TTTCTTTCTGATTTTCTTTCTTTGATTTAGGCTGATGAGCTGATGAGCTGATGAG 3029
 Db 534 TTTCTTTCTGATTTTCTTTCTTTGATTTAGGCTGATGAGCTGATGAGCTGATGAG 475
 Qy 3030 CACTTCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3089
 Db 474 CACTTCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 415
 Qy 3090 AGTCTGCTTTTGGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3149
 Db 414 AGTCTGCTTTTGGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 370
 Qy 3150 TCTACATTCATCAATGTTGTGAAACCTGTTGACATGATTTATGTTGCTGCTGTT 3209
 Db 369 ----- 370
 Qy 3210 TGATTTCTGTTATTTATGATTTTGGATGAGGGAATGCTGAGATTTGATGATGATGATG 3269
 Db 369 -----AGGTTTGGATGAGGGAATGCTGAGATTTGATGATGATGATGATGATGATG 326
 Qy 3270 GCTGAACTGCGCAGCTTGGGTTGTTTATGATTTATGATTTATGATTTATGATTTATG 3329
 Db 325 GCTGAACTGCGCAGCTTGGGTTGTTTATGATTTATGATTTATGATTTATGATTTATG 266
 Qy 3330 AGTGTGACAGTCTGACGATGAGAACCCGCTGCTGTTGTTGTTGTTGTTGTTGTTGTT 3389
 Db 265 AGTGTGACAGTCTGACGATGAGAACCCGCTGCTGTTGTTGTTGTTGTTGTTGTTGTT 206
 Qy 3390 GCTGTCTATGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 3449
 Db 205 GCTGTCTATGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 146
 Qy 3450 ACAAGATGCAAGTTTCTGTTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 3509
 Db 145 ACAAGATGCAAGTTTCTGTTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 86
 Qy 3510 GCATCATATATATCTCATATATGATGATGATGATGATGATGATGATGATGATGATGAT 3569
 Db 85 GCATCATATATATCTCATATATGATGATGATGATGATGATGATGATGATGATGATGAT 26
 Qy 3570 ATTCTCTCTTAAGTTAGTTTAT 3594
 Db 25 ATTCTCTCTTAAGTTAGTTTAT 1
 RESULT 9
 ADJ38264
 ID ADJ38264 standard; cDNA; 631 BP.
 AC ADJ38264;
 XX
 DT 06-MAY-2004 (first entry)
 DE Plastid division-related Arc6 orthologue cDNA 51.
 XX
 KW prokaryotic type; plastid division; Ftn2; ARC6; ARCS; Fzo; plant cell;

DB 262 TGTTCATATCCGGAGAACATGCTCTAAATGTAATCTTACACACAGATATGAGAT 321
QY 3464 TTTCGGTCCAGTAGGGTGGAAATCAGTCGAGGCTGTCTTGATCATATATAC 3523
DB 322 GTCTTGTTCACACTGAGGCTGGAAATCAGTAGAGATCTGTCTACAAATCTTAATG 381
QY 3524 TCATATGTAGCAT 3536
DB 382 ATGTATAAGCAT 394

RESULT 15
ACN48855/C
ID ACN48855 standard; cDNA; 552 BP.
XX
AC ACN48855;
XX
DT 02-DEC-2004 (first entry)
XX
DE Cotton primed seed EST Clone ID: LIB3825-027-Q6-N6-H1, SEQ:3636.
XX
KM Cotton, plant; EST; expressed sequence tag; transgenic plant; seed;
KM variety DP50B; library LIB3825; molecular tag; molecular marker;
KM genetic mapping; molecular mapping; seed germination; plant growth;
KM plant quality; plant yield; plant breeding; tissue printing; ss.
XX
OS Gossypium hirsutum.
XX
PN US2004123340-A1.
XX
PD 24-JUN-2004.
XX
PF 12-DEC-2001; 2001US-00021323.
XX
PR 14-DEC-2000; 2000US-0255619P.
XX
PA (DEIK/) DEIKMAN J.
PA (FENG/) FENG P C C.
PA (FINCH/) FINCHER K L.
PA (ZIEGL/) ZIEGLER T E.
XX
PI Deikman J, Feng PCC, Fincher KL, Ziegler TE;
XX
DR WPI; 2004-479808/45.
XX
PT New isolated nucleic acid molecule that encodes a plant protein or its
PT fragment, useful for isolating a variety of agronomically significant
PT genes associated with plant growth, quality or yield, and as molecular
PT tags to map genes.
XX
PS Claim 1; SEQ ID NO 3636; 34bp; English.
XX
CC The invention relates to 17880 cotton expressed sequence tags (ESTs;
CC ACN45220-ACN63099). The ESTs were isolated from cDNA libraries generated
CC from primed or non-primed seeds from variety DP50B, mature seeds from
CC variety Coker 312 Boswell 96 field, and androecium tissue, gynoecium
CC tissue, developing fibres, carpel walls and septa from variety
CC Nuoton33B. The invention also relates to substantially purified
CC proteins or their fragments encoded by nucleic acid molecules of the
CC invention, and to transformed plants having a nucleic acid construct
CC comprising a nucleic acid of the invention. The cotton ESTs are useful as
CC molecular tags to isolate genetic regions, to isolate genes, to map
CC genes, to determine gene function and to determining whether genes are
CC members of a particular gene family. The nucleic acid molecules may be
CC used for isolating a variety of agronomically significant genes
CC associated with plant growth, quality, yield, and could also serve as
CC links in metabolic and catabolic pathways. The nucleic acid molecules are
CC also useful for identifying genes important in initiating and maintaining
CC seed germination or that may be used to mitigate stresses encountered
CC during seed germination. The ESTs additionally enable the acquisition of
CC promoters and cis-regulatory elements which will be useful to express
CC agronomically significant genes in these tissues and/or other tissues,

CC and also permits the acquisition of molecular markers useful in breeding
CC schemes, genetic and molecular mapping, and in cloning of agronomically
CC significant genes. The nucleic acid molecules are further useful for
CC detecting the expression level or pattern of a protein or mRNA and for
CC detecting the presence or quantity of a protein by tissue printing. The
CC present sequence represents a specifically claimed EST isolated from a
CC cotton variety DP50B primed seed cDNA library (LIB3825). The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from the US patent office at
CC seqdata.uspco.gov/sequence.html?docid=US20040123340
XX

SO Sequence 552 BP; 153 A; 122 C; 110 G; 167 T; 0 U; 0 Other;

Query Match 4.3%; Score 159.4; DB 13; length 552;

Best Local Similarity 69.3%; Pred. No. 1.2e-30;

Matches 217; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 3224 ATAGCTTTGGATGGCGGAATGCTGAAGATTGGACTGACAGACGACTGAAGTCGCCA 3283
DB 444 AGAGGTTCTGGATGGTCAAAATGTTGAAGACATGACAGATCGACCCGAAATCGCTCA 385
QY 3284 GCTTGGGTTGGTTTANGATTATACACTGTTGAACATATCTGTTGACAGTGAACAGTCTC 3343
DB 384 GCTTGGTGGGTATATAATATAGTCTACTGACATGCGCATTGACAGTGTACCTTTTC 325
QY 3344 AGCAGATGAACCCGTCCTCTGTGGAAGCACTCTGGAAGAGTCTGCTTATCTGA 3403
DB 324 ACTAGATGGCCAGGAGCTGTAGTGAAGCTACTCTGGAAGAAATCCACTGCTTGACTGA 265
QY 3404 TTTCGTTATCCAGAAAACAATGCTACTGATGTGCAAGACTTACACACAACAATACGAAGT 3463
DB 264 TGTTCATCATCCCGAGAAACAATGCTCTTAATGTGMAACTCTTACACCAACGAGATATGAGT 205
QY 3464 TTTCGGTCCAGGTCAGAGGTGAAAAATCAGTGAAGGCTGTCTTGATCATATATATAC 3523
DB 204 GTCTTGTTCACACTCAGGCTGGAATAATCACTGAAGGATCTGTCTACAAATCTTAGCTATG 145
QY 3524 TCATATGTAGCAT 3536
DB 144 ATGTATAAGCAT 132

Search completed: December 10, 2005, 19:22:07
Job time : 2104.73 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 10, 2005, 16:57:24 ; Search time 18153.9 Seconds
(without alignments)
11482.131 Million cell updates/sec

Title: US-10-600-070B-3

Perfect score: 3667
1 tgcctgcattcaaggagaaat.....ctataacataaagggtacaa 3667

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenBdb1:*
1: gb_da:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_srs:*
11: gb_sy:*
12: gb_un:*
13: gb_vl:*
14: gb_hcg:*
15: gb_pl:*

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3664	99.9	3668	15	AY221468 Arabidops
2	3660.8	99.8	3664	15	AY221467 Arabidops
3	3659	99.8	85791	15	AB016888 Arabidops
4	1931	52.7	2637	15	AY091075 Arabidops
5	1752.8	47.8	2438	15	AY221469 Arabidops
6	1744.6	47.6	2436	15	AY150490 Arabidops
7	338.6	9.2	133779	14	AC158210 Medicago
8	317.6	8.7	110000	15	AP008208_011
9	317.6	8.7	150462	15	AP004885 Oryza sat
10	246.8	6.7	184752	14	AC155597 Zea mays
11	170.4	4.6	550	10	D0022571 Beta Vulg
12	146.2	4.0	117818	14	AC160013 Medicago
13	57.6	1.6	7218	6	I66494
14	54	1.5	1141	6	ARS79680
15	54	1.5	1141	6	AX083744
16	54	1.5	250029	2	AE014820
17	53.8	1.5	205796	14	AL713891 Mus muscu
18	53.8	1.5	215524	9	AL603913 Mouse DNA

19	53.8	1.5	250078	2	AE014829	AE014829 Plasmodi
20	53.8	1.5	333321	2	AC116986	AC116986 Dictyoste
21	53.2	1.5	198573	14	AC110118	AC110118 Rattus no
22	53	1.4	184039	9	AC139376	AC139376 Mus muscu
23	53	1.4	232373	14	AC134805	AC134805 Rattus no
24	52.8	1.4	617	10	BV331985	BV331985 S230P6420
25	52.6	1.4	222366	14	AC119799	AC119799 Pan trogl
26	52.4	1.4	234008	14	CR392350	CR392350 Dabo rer
27	52.4	1.4	264451	14	AC158594	AC158594 Mus muscu
28	52.2	1.4	20079	1	AE017244_09	Continuation (10 o
29	52.2	1.4	97405	1	AE017243_8	Continuation (9 of
30	52.2	1.4	110000	1	AE017244_08	Continuation (9 of
31	52	1.4	180084	9	AC140372	AC140372 Mus muscu
32	52	1.4	245924	5	BX004773	BX004773 Zebrafish
33	52	1.4	255036	9	AC118206	AC118206 Mus muscu
34	52	1.4	268597	14	AC094787	AC094787 Rattus no
35	51.6	1.4	2000	6	AX655393	AX655393 Sequence
36	51.6	1.4	23213	8	AL590436	AL590436 Human DNA
37	51.6	1.4	110000	14	AC108609_2	Continuation (3 of
38	51.6	1.4	292533	14	AC107260	AC107260 Rattus no
39	51.4	1.4	135256	9	AC163613	AC163613 Mus muscu
40	51.4	1.4	139995	9	AC131797	AC131797 Mus muscu
41	51.4	1.4	150579	9	AC126937	AC126937 Mus muscu
42	51.4	1.4	171307	8	AC104020	AC104020 Homo sapi
43	51.4	1.4	193634	14	AC068495	AC068495 Mus muscu
44	51.4	1.4	249854	14	CT009658	CT009658 Mus muscu
45	51.4	1.4	253094	14	AC113774	AC113774 Rattus no

ALIGNMENTS

RESULT 1	AY221468	3668 bp	DNA	linear	PLN 05-AUG-2003
LOCUS	AY221468				
DEFINITION	Arabidopsis thaliana division protein (ARCF) gene, complete cds.				
ACCESSION	AY221468				
VERSION	AY221468.1	GI:33436338			
KEYWORDS					
SOURCE	Arabidopsis thaliana (thale cress)				
ORGANISM	Arabidopsis thaliana				
REFERENCE	1 (bases 1 to 3668)				
AUTHORS	Vittha,S., Froehlich,J.F., Kokeharova,O., Pyke,K.A., Van Ert,H. and Oosteryoung,K.W.				
TITLE	ARCF is a J-Domain Placitid Division Protein and an Evolutionary				
JOURNAL	Descendant of the Cyanobacterial Cell Division Protein Ftn2				
PUBMED	Plant Cell 15 (8), 1918-1933 (2003)				
REFERENCES	12897262				
AUTHORS	2 (bases 1 to 3668)				
TITLE	Vittha,S., Kokeharova,O., van Ert,H., Froehlich,J.E. and Oosteryoung,K.W.				
JOURNAL	Direct Submission				
PUBMED	Submitted (17-JAN-2003) Department of Plant Biology, Michigan State				
REFERENCES	University, 166 Plant Biology Building, East Lansing, MI 48824, USA				
AUTHORS	Location/Qualifiers				
TITLE	1. 3668				
JOURNAL	/organism="Arabidopsis thaliana"				
FEATURES	/mol type="genomic DNA"				
SOURCE	/db xref="taxon:3702"				
	/ecotype="Wassilewskija"				
gene	478..3515				
	/gene="ARCF"				
	/note="synonym: At5g42480"				
mRNA	join(<478..984,1067..1291,1379..1639,1723..2710,				
	3000..3133,3225..3515)				
	/gene="ARCF"				
	/product="division protein"				
CDS	join(478..984,1067..1291,1379..1639,1723..2710,3000..3133,				
	3225..3515)				

```
/gene="ARC6"  
/note="inner envelope membrane-localized; similar to  
cyanobacterial cell division protein sli0169 of  
Synchocystis; contains J-domain"  
/codon_start=1  
/product="division protein"  
/protein_id="AA018645.1"  
/db_xref="GI:33436339"  
/translation="MEALSHVIGIGLSPQLRLPPTTKLRSHNTSTTICASKMAD  
RLSDPNTSDSSSSFPATTTATTAIVLPSIDRBEHVLPIDFVOVLAOYHPLT  
DTRAPAPVSKPQPGSPDALSIRPOLIQAACETLSREYVGLDDEBPT  
VITDPMRKVPGLCVLOEGSTELVLRVGLAKERLPKSPKQDVUVALALDYS  
RDAAALDPDPITIGBEFVEEALKLQEBGASLAPDLQAQIDLEBETPRVLELLS  
LPLGDDYAKRLNGLSGVRNLIMVSGGASALVGLTRKFNMEAFLRMTAAQVLD  
FVATPSNIAPASFVEVALVAQAFIKGPHILODADQFOOLQOKVMAEIPAM  
LYDRNWEIDFGEERGLCALLIGKDECRMLIGLSDSDQYRMLVLEFVLENSND  
DNDDLRLGLKLETMLAGVFRPRDPTDKKPKLGDYDDPMVLSYLRVVGSPPL  
AAATMARIGAEHYKASAMOLQKVPSTRYDRNSAEKDVQETVFSVDPRVNGRD  
GEPEVFAEAVRSENFETNDYALRAVSESSVDEETVEMSVADMLKEASVKILAAQV  
AIGLISLPSQRYPLKSSSFKDKMVSMSDVATIGSRADDEBALPRMDARTAEI  
VSKQIKSLAFGPDIRIEMLPVLDGMLKIWDRAEYRQQLGVLDYTLKLSVS  
VTVSADGTRALVEATLEESACLSDLVHEENNATDVRYTTRVEVFWSSGKITEGSV  
LAS"
```

ORIGIN

Query Match 99.9%; Score 3664; DB 15; Length 3668;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 3664; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
OY 4 TCTGCATTAAAGCAATACATTATTAAGCAATTTGCTGATTTCAACAAGTTTGGTT 63  
Db 1 TCTGCATTAAAGCAATACATTATTAAGCAATTTGCTGATTTCAACAAGTTTGGTT 60  
OY 64 GGCCTATAGATTCAATGGCTCTGTTGCTTTTACATTTACATGTCATAATAGTTTGCAT 123  
Db 61 GGCCTATAGATTCAATGGCTCTGTTGCTTTTACATTTACATGTCATAATAGTTTGCAT 120  
OY 124 TTTCACATTTTCAGTTGATGTTAAAGAAAGAGAGGAATTGATGGGGTTTGTGGGTTT 183  
Db 121 TTTCACATTTTCAGTTGATGTTAAAGAAAGAGAGGAATTGATGGGGTTTGTGGGTTT 180  
OY 184 AAACCTTAAAGTAGCAAGAAATTAAAGTATGTTTACGTTGCTCTATATGTGTAAT 243  
Db 181 AAACCTTAAAGTAGCAAGAAATTAAAGTATGTTTACGTTGCTCTATATGTGTAAT 240  
OY 244 GAAGGCACTCCAAAGGTTCTTAGGTGAATAGATTATTTAGACATTTAATCATATAA 303  
Db 241 GAAGGCACTCCAAAGGTTCTTAGGTGAATAGATTATTTAGACATTTAATCATATAA 300  
OY 304 GTCCGTGGCACTGTAAATCATAGATTGTTTTTATTTTTTTCAGTAGCTGTGATGTT 363  
Db 301 GTCCGTGGCACTGTAAATCATAGATTGTTTTTATTTTTTTCAGTAGCTGTGATGTT 360  
OY 364 TTTTGATTTAACTTATCTAATCAAAATCCATAAACCCTTACAGCAACAAAG 423  
Db 361 TTTTGATTTAACTTATCTAATCAAAATCCATAAACCCTTACAGCAACAAAG 420  
OY 424 TCTCTTCAATATGTAAACAGAAAGTTTGTGATAGCTTAAAGAAAGCACTCCCATG 483  
Db 421 TCTCTTCAATATGTAAACAGAAAGTTTGTGATAGCTTAAAGAAAGCACTCCCATG 480  
OY 484 GAAGCTTGAAGTCAAGTGGGATGGTCTCTCCCATTCGAATTAAGCGATTACACAG 543  
Db 481 GAAGCTTGAAGTCAAGTGGGATGGTCTCTCCCATTCGAATTAAGCGATTACACAG 540  
OY 544 GCGAGCAAAAGCTCCGAGTAGCAACACCTTACAACTATCTGCTCCGACAGCAA 603  
Db 541 GCGAGCAAAAGCTCCGAGTAGCAACACCTTACAACTATCTGCTCCGACAGCAA 600  
OY 604 TGGGCGGACCGCTTCTCCGAGCTTCAATTTACCTCGAATCTCTCTCCGCTCTC 663  
Db 601 TGGGCGGACCGCTTCTCTCGAATTTCACCTCGAATCTCTCTCTCTCTCTC 660
```

```
OY 664 GCCACCGCACCAACACCGGCACTGCTCTCTGACCATCTATGATCGTCCGAA 723  
Db 661 GCCACCGCACCAACACCGGCACTGCTCTCTGACCATCTATGATCGTCCGAA 720  
OY 724 CGCCAGCTCCCATCCCATTTGATTTTACAGGATTTAGAGCTGAAACATTTCTTA 783  
Db 721 CGCCAGCTCCCATCCCATTTGATTTTACAGGATTTAGAGCTGAAACATTTCTTA 780  
OY 784 ACCGATGAATTCAGAAAGATTCGAAGCTAAGGTTTCGAAACCGCGCAATTCGTTTC 843  
Db 781 ACCGATGAATTCAGAAAGATTCGAAGCTAAGGTTTCGAAACCGCGCAATTCGTTTC 840  
OY 844 AGCAGCAACGCTTAACTACCGCGAGACAGATTTCTCAAGCTGCTCGAAACTCTCT 903  
Db 841 AGCAGCAACGCTTAACTACCGCGAGACAGATTTCTCAAGCTGCTCGAAACTCTCTCT 900  
OY 904 AATCTCGGCTCTGAAAGAGTCAAGAAAGCTCTTGTATGATGAAGAAAGCTACAGTC 963  
Db 901 AATCTCGGCTCTGAAAGAGTCAAGAAAGCTCTTGTATGATGAAGAAAGCTACAGTC 960  
OY 964 ATCACTAGTCTCTTGGGATTAAGTAAATTTGATTTCCGAAATTAATTAAGTTTCTTCT 1023  
Db 961 ATCACTAGTCTCTTGGGATTAAGTAAATTTGATTTCCGAAATTAATTAAGTTTCTTCT 1020  
OY 1024 TTAATTTCAATGAATGATTAAGAGAGAACTTTTATCTAGTGAAGGTTCTTGGGCTCT 1083  
Db 1021 TTAATTTCAATGAATGATTAAGAGAGAACTTTTATCTAGTGAAGGTTCTTGGGCTCT 1080  
OY 1084 CTGCTATTCGAAAGAGTGTGAGACTGATGATGTTCTCGGTTGCTGAGGCTCTGCT 1143  
Db 1081 CTGCTATTCGAAAGAGTGTGAGACTGATGATGTTCTCGGTTGCTGAGGCTCTGCT 1140  
OY 1144 TTAAGAGAGGTTGCTAAGTCGTTTAAAGAAAGTGTGTTTAAAGTGTGCTTGGCT 1203  
Db 1141 TTAAGAGAGGTTGCTAAGTCGTTTAAAGAAAGTGTGTTTAAAGTGTGCTTGGCT 1200  
OY 1204 TCTCGATGTCGAGGAGTCTATGCGATGATTCACCTGATTTTATCTGCTTATGA 1263  
Db 1201 TCTCGATGTCGAGGAGTCTATGCGATGATTCACCTGATTTTATCTGCTTATGA 1260  
OY 1264 GTTGTGTAAGAAAGCTTTGAAGCTTTTACAGATGTTGACTGCTTGGTAAATTTGAGC 1323  
Db 1261 GTTGTGTAAGAAAGCTTTGAAGCTTTTACAGATGTTGACTGCTTGGTAAATTTGAGC 1320  
OY 1324 AGCGTGGCTTATTAAGAACTTTCTTGATTTGATCTTTGATTTGAGTCTTGTGAGA 1383  
Db 1321 AGCGTGGCTTATTAAGAACTTTCTTGATTTGATCTTTGATTTGAGTCTTGTGAGA 1380  
OY 1384 GGAAGAGCAAGTAGCTTGCACCGGATTTACGTGCACAAATTGATGAGACTTTGAGA 1443  
Db 1381 GGAAGAGCAAGTAGCTTGCACCGGATTTACGTGCACAAATTGATGAGACTTTGAGA 1440  
OY 1444 GATGACTCCGCGTATGTCCTTGAGAGTACTTGGCTTACCGCTTGGTATTAAGCTGC 1503  
Db 1441 GATGACTCCGCGTATGTCCTTGAGAGTACTTGGCTTACCGCTTGGTATTAAGCTGC 1500  
OY 1504 GAAAAGACTAATGTTTAAAGCGGTGCGGAAATATTTGTGCTGTGTTGAGAGGTG 1563  
Db 1501 GAAAAGACTAATGTTTAAAGCGGTGCGGAAATATTTGTGCTGTGTTGAGAGGTG 1560  
OY 1564 AGCATGAGCTCTTGTGGGGGTTTGAACCGGTGAGAAATTATGAATGAGCGTTTTTACG 1623  
Db 1561 AGCATGAGCTCTTGTGGGGGTTTGAACCGGTGAGAAATTATGAATGAGCGTTTTTACG 1620  
OY 1624 AATGACAGCTCTGAGAGGATTAACGTTTATGATCTTTTATTTTATTTTATGAGTGA 1683  
Db 1621 AATGACAGCTCTGAGAGGATTAACGTTTATGATCTTTTATTTTATTTTATGAGTGA 1680  
OY 1684 TATPACTTATGATTTTCAATTTAATGATGTTGTGAGTGAATCTTTTGTGAGT 1743  
Db 1681 TATPACTTATGATTTTCAATTTAATGATGTTGTGAGTGAATCTTTTGTGAGT 1740  
OY 1744 ACCCAAGCAATATTCAGAGAGTCAATTGAAGTTTACGAAGTTGCACTTGTCTTGTG 1803
```

Db 1741 ACCCAAGCAATATCTCAGAGAGTCAATTTGAGATTCAGAACTTCACCTGCTCTG 1800
Qy 1804 GCTCAAGCTTTTATGTAAGAACCAACCTTTACAGGATGCTGATAGCAATTCAG 1863
Db 1801 GCTCAAGCTTTTATGTAAGAACCAACCTTTACAGGATGCTGATAGCAATTCAG 1860
Qy 1864 CAACCTCAGAGGCTAAGGTAATGCTATGAGATTCCTGCGATGTTGATAGTACAGG 1923
Db 1861 CAACCTCAGAGGCTAAGGTAATGCTATGAGATTCCTGCGATGTTGATAGTACAGG 1920
Qy 1924 AATTAATGGAGATTAACCTTCGCTCTAGAAAGGGGACTCTGTCACCTGCTATAGGCAA 1983
Db 1921 AATTAATGGAGATTAACCTTCGCTCTAGAAAGGGGACTCTGTCACCTGCTATAGGCAA 1980
Qy 1984 GTTATGAAATCCGATGATGCTGAGGCTTACAGATGAGGATTCACAAATATAGGAATCA 2043
Db 1981 GTTATGAAATCCGATGATGCTGAGGCTTACAGATGAGGATTCACAAATATAGGAATCA 2040
Qy 2044 GCTATTTGAGAGTTTGTGTAAGAAATTCGATGATGATGATGATGATGATGATGATG 2103
Db 2041 GCTATTTGAGAGTTTGTGTAAGAAATTCGATGATGATGATGATGATGATGATGATG 2100
Qy 2104 CTATGCAATTTGTTGAAACCTGCTGAGGAGGCTGCTCTGCTGCTGCTGCTGCTGCTG 2163
Db 2101 CTATGCAATTTGTTGAAACCTGCTGAGGAGGCTGCTCTGCTGCTGCTGCTGCTGCTG 2160
Qy 2164 AATGATTAATTAATTTAACTCGGAGGACTATGATGATGATGATGATGATGATGATGATG 2223
Db 2161 AATGATTAATTAATTTAACTCGGAGGACTATGATGATGATGATGATGATGATGATGATG 2220
Qy 2224 GAAAGATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2283
Db 2221 GAAAGATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2280
Qy 2284 GAGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2343
Db 2281 GAGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2340
Qy 2344 TATACAGATGAGAACTCGGCTGAAACCAAGGATGTCAGAGAGAGAGGATGATGATGATG 2403
Db 2341 TATACAGATGAGAACTCGGCTGAAACCAAGGATGTCAGAGAGAGAGGATGATGATGATGATG 2400
Qy 2404 CCTGTTGTAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2463
Db 2401 CCTGTTGTAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2460
Qy 2464 AAGACCTCTGAAACCTTTGAAACTATGATGATGATGATGATGATGATGATGATGATGATGATG 2523
Db 2461 AAGACCTCTGAAACCTTTGAAACTATGATGATGATGATGATGATGATGATGATGATGATGATG 2520
Qy 2524 AGCGTTGATGAAACTCTGTTGAAATGTCGTTGCTGATGATGATGATGATGATGATGATGATG 2583
Db 2521 AGCGTTGATGAAACTCTGTTGAAATGTCGTTGCTGATGATGATGATGATGATGATGATGATG 2580
Qy 2584 AAGATCTTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2643
Db 2581 AAGATCTTAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2640
Qy 2644 CTTTAAAGCAGCTCATCTTTTCAACGAGATGATGATGATGATGATGATGATGATGATGATG 2703
Db 2641 CTTTAAAGCAGCTCATCTTTTCAACGAGATGATGATGATGATGATGATGATGATGATGATG 2700
Qy 2704 GCTACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2763
Db 2701 GCTACCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2760
Qy 2764 TTGTTTGTAGCTAGAGACATAGTCCACTTAATATAGATGATGATGATGATGATGATGATG 2823
Db 2761 TTGTTTGTAGCTAGAGACATAGTCCACTTAATATAGATGATGATGATGATGATGATGATG 2820
Qy 2824 ATTAACAGTGTGCTAGTAAATTTCTCATATATGCTGCTGATGATGATGATGATGATGATG 2883

Db 2821 ATTAACAGTGTGCTAGTAAATTTCTCATATATGCTGCTGATGATGATGATGATGATGATG 2880
Qy 2884 TAGACAGAAATGTAATTTCACTCTCAACATTTCTGTTAGAAATGATGATGATGATGATGATG 2943
Db 2881 TAGACAGAAATGTAATTTCACTCTCAACATTTCTGTTAGAAATGATGATGATGATGATGATG 2940
Qy 2944 TTGCTTAGTGTGCTTTGTCACACTTTTCTTTCTGATGATGATGATGATGATGATGATGATG 3003
Db 2941 TTGCTTAGTGTGCTTTGTCACACTTTTCTTTCTGATGATGATGATGATGATGATGATGATG 3000
Qy 3004 GTACATGAGAGCTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3063
Db 3001 GTACATGAGAGCTGACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3060
Qy 3064 TATAGTATCAAGTGGCAGAAATTAAGTCTGCTGCTTTGAGGCTGATACCGCATAGA 3123
Db 3061 TATAGTATCAAGTGGCAGAAATTAAGTCTGCTGCTTTGAGGCTGATACCGCATAGA 3120
Qy 3124 AATGTTACAGAGGATGAGGAAATTAATCTACATTCATTCATTCATTCATTCATTCATTCATTC 3183
Db 3121 AATGTTACAGAGGATGAGGAAATTAATCTACATTCATTCATTCATTCATTCATTCATTCATTC 3180
Qy 3184 ACATGATTTATGATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3243
Db 3181 ACATGATTTATGATCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3240
Qy 3244 TGCTGAAGATTTGACCTGACAGAGAGCTGAACTGCGCAGCTGCTGCTGCTGCTGCTGCTGCTG 3303
Db 3241 TGCTGAAGATTTGACCTGACAGAGAGCTGAACTGCGCAGCTGCTGCTGCTGCTGCTGCTGCTG 3300
Qy 3304 ATACATGTTGAACTATCTGTTGACAGATGATGATGATGATGATGATGATGATGATGATGATGATG 3363
Db 3301 ATACATGTTGAACTATCTGTTGACAGATGATGATGATGATGATGATGATGATGATGATGATGATG 3360
Qy 3364 TGTGGAAGCAACTCTGGAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3423
Db 3361 TGTGGAAGCAACTCTGGAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3420
Qy 3424 ATGCTACTGATGTCGAGAACTACACAAAGATGATGATGATGATGATGATGATGATGATGATGATG 3483
Db 3421 ATGCTACTGATGTCGAGAACTACACAAAGATGATGATGATGATGATGATGATGATGATGATGATG 3480
Qy 3484 GGAATATCACTGAAGGCTCTGTTCTTGCAATATATATATATATATATATATATATATATATAT 3543
Db 3481 GGAATATCACTGAAGGCTCTGTTCTTGCAATATATATATATATATATATATATATATATATAT 3540
Qy 3544 CTGCGAATTTCTCTTTGTTCTGTAATTTCTCTCTAAGTTAGTTTATTAATGAACA 3603
Db 3541 CTGCGAATTTCTCTTTGTTCTGTAATTTCTCTCTAAGTTAGTTTATTAATGAACA 3600
Qy 3604 CAAAAAATTAACGTTCTTGACACACCTTTCTGATCTAATCTAATCAATTAAGGCT 3663
Db 3601 CAAAAAATTAACGTTCTTGACACACCTTTCTGATCTAATCTAATCAATTAAGGCT 3660
Qy 3664 ACAA 3667
Db 3661 ACAA 3664

RESULT 2
AY221467
LOCUS AY221467 3664 bp DNA linear PLN 05-AUG-2003
DEFINITION Arabidopsis thaliana truncated division protein (arce6) gene, arce6-1
allele, complete cds; nuclear gene for chloroplast product.
ACCESSION AY221467
VERSION AY221467.1 GI:3436274
KEYWORDS
SOURCE Arabidopsis thaliana (chale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE
1 (bases 1 to 3664)

AUTHORS Vitha, S., Froehlich, J.E., Koksharova, O., Pyke, K.A., Van Erp, H. and Osereroung, K.W.
TITLE ARC6 is a J-Domain Plastid Division Protein and an Evolutionary Descendant of the Cyanobacterial Cell Division Protein Fnz2
JOURNAL Plant Cell 15 (8), 1918-1933 (2003)
PUBMED 12897262
2 (bases 1 to 3664)
Vitha, S., Koksharova, O., van Erp, H., Froehlich, J.E. and Osereroung, K.W.
TITLE Direct Submission
JOURNAL Submitted (17-JAN-2003) Department of Plant Biology, Michigan State University, 166 Plant Biology Building, East Lansing, MI 48824, USA
FEATURES
source
1..3664
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/ecotype="Ws-2"
/note="genotype: arc6 mutant"
1..3664
/gene="arc6"
/allele="arc6-1"
join(<478..984,1067..1291,1379..>1621)
/gene="arc6"
/allele="arc6-1"
/product="truncated division protein"
join(478..984,1067..1291,1379..1621)
/gene="arc6"
/allele="arc6-1"
/note="truncation caused by premature stop codon"
/codon_start=1
/product="truncated division protein"
/protein_id="AAQ18644.1"
/db_xref="GI:33436275"
/ORIGIN
Query Match 99.8%; Score 3660.8; DB 15; Length 3664;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3662; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
4 TCTGATTAGAGAAATCAATTATTAAGCAATTGCTTGAATTAACAAGATTTGGCT 63
1 TCTGATTAGAGAAATCAATTATTAAGCAATTGCTTGAATTAACAAGATTTGGCT 60
64 GGCATTAGAGATTCATTGGCTCTGTTTGCTTTACATTTACATGTCATTAATAGTTGCAAT 123
61 GGCATTAGAGATTCATTGGCTCTGTTTGCTTTACATTTACATGTCATTAATAGTTGCAAT 120
124 TTAAACATTTCAGTTGGATGTTAAAGAAAAGAGGAAATGATGGGCTTTTGGGCTTT 183
121 TTAAACATTTCAGTTGGATGTTAAAGAAAAGAGGAAATGATGGGCTTTTGGGCTTT 180
184 AAACCTTAAAGTAGCAAGATTAAGTCAATGGTTTACGTTGCTCTATATGTTGTTAAAT 243
181 AAACCTTAAAGTAGCAAGATTAAGTCAATGGTTTACGTTGCTCTATATGTTGTTAAAT 240
244 GAAGCAACTCCACGGTTCTTAGTGGAATGATTAATTTAGACGATTTAAACATCATAA 303
241 GAAGCAACTCCACGGTTCTTAGTGGAATGATTAATTTAGACGATTTAAACATCATAA 300
304 GTCCGGTGGAGCTGTAACATCATAGATTGTTTTTATTTTTCAGTAGCTGTGATGTT 363
301 GTCCGGTGGAGCTGTAACATCATAGATTGTTTTTATTTTTCAGTAGCTGTGATGTT 360
364 TTTTGAATTAACCTTAACACTCAAAATCAAAATCCATAAACCCTAGACGACCAACG 423
361 TTTTGAATTAACCTTAACACTCAAAATCAAAATCCATAAACCCTAGACGACCAACG 420

424 TCTCTTCAATATATGTAATAAACAAGAACTTTTGTAGTACCTTAATAAAGACACTCCATG 483
421 TCTCTTCAATATATGTAATAAACAAGAACTTTTGTAGTACCTTAATAAAGACACTCCATG 480
484 GAAGCTTGAGTCAAGCTGCGGATTTGGTCTCTCCCATTTCAATTAATAGCCATTAACACG 543
481 GAAGCTTGAGTCAAGCTGCGGATTTGGTCTCTCCCATTTCAATTAATAGCCATTAACACG 540
544 GCGACGCAAAAGCTCCGAGTAGGCAACACCTCTCAACTATCTGCTCCGACAGAAA 603
541 GCGACGCAAAAGCTCCGAGTAGGCAACACCTCTCAACTATCTGCTCCGACAGAAA 600
604 TGGGCGACCGCTCTTCTCTCCGACTTCAATTTCACTCCGATTTCTGCTCTCTCTTC 663
601 TGGGCGACCGCTCTTCTCTCCGACTTCAATTTCACTCCGATTTCTGCTCTCTCTTC 660
664 GCGACCGGCAACACACCGGCACTCTGCTCTCTGCAACCATCTATTGATTCGTCGCGAA 723
661 GCGACCGGCAACACACCGGCACTCTGCTCTCTGCAACCATCTATTGATTCGTCGCGAA 720
724 GCGACGCTCCCATCCCATTTGATTTCTACCAAGTATTAAGAGCTCAAAACATTTCTTA 783
721 GCGACGCTCCCATCCCATTTGATTTCTACCAAGTATTAAGAGCTCAAAACATTTCTTA 780
784 ACCGATGGAATCAGAAAGCAATTCGAAAGCTAGGCTTTGAAACCGCGCAATTCGTTTC 843
781 ACCGATGGAATCAGAAAGCAATTCGAAAGCTAGGCTTTGAAACCGCGCAATTCGTTTC 840
844 AGCGACGCGCTTTAATCAACCGGAGACATATCTTCAACCTGCTCCGAACTCTGCT 903
841 AGCGACGCGCTTTAATCAACCGGAGACATATCTTCAACCTGCTCCGAACTCTGCT 900
904 AATCTCGCTCTAGAAAGAGATCAATGAAGTCTTCTGATGATGAAGAGCTACAGTC 963
901 AATCTCGCTCTAGAAAGAGATCAATGAAGTCTTCTGATGATGAAGAGCTACAGTC 960
964 ATCACTGATGTTCTTGGAGTAAGTAAATTTGATTTGCGAATTAATAAAGTTTCTGCT 1023
961 ATCACTGATGTTCTTGGAGTAAGTAAATTTGATTTGCGAATTAATAAAGTTTCTGCT 1020
1024 TTAATTTCAATGAAATGGAATAAAGAAAGCACTTATCTATGTAAGGTTCTCGGGCTCT 1083
1021 TTAATTTCAATGAAATGGAATAAAGAAAGCACTTATCTATGTAAGGTTCTCGGGCTCT 1080
1084 CTGTGATGTAAGAAAGTGGTGAAGTGAATAGTTCTTCTCGGTTGGTGAAGCTCTGCT 1143
1081 CTGTGATGTAAGAAAGTGGTGAAGTGAATAGTTCTTCTCGGTTGGTGAAGCTCTGCT 1140
1144 TTAAGAAAGTGGCTTAAGTGGTTTAAACAAGATGTTGTTTAAATGAGGCTTGCCTT 1203
1141 TTAAGAAAGTGGCTTAAGTGGTTTAAACAAGATGTTGTTTAAATGAGGCTTGCCTT 1200
1204 TCTGATGTCCTCGAGGATGCTATGGAATGGAATCCACCTGATTTTATTTACGTTATGA 1263
1201 TCTGATGTCCTCGAGGATGCTATGGAATGGAATCCACCTGATTTTATTTACGTTATGA 1260
1264 GTTGTGAGGAAAGCTTTGAAGCTTTTACAGTATGTTGAAGTCTTGTGTAATTTGACG 1323
1261 GTTGTGAGGAAAGCTTTGAAGCTTTTACAGTATGTTGAAGTCTTGTGTAATTTGACG 1320
1324 AGCGTTGGCTTAAAGAACTTTTGAAGCTTTTACAGTATGTTGAAGTCTTGTGTAATTTG 1383
1321 AGCGTTGGCTTAAAGAACTTTTGAAGCTTTTACAGTATGTTGAAGTCTTGTGTAATTTG 1380
1384 GGAAGGAGCAAGTACCTTGACCGGATTAACGAGCAAAATGATAGAGACTTTGAGAGA 1443
1381 GGAAGGAGCAAGTACCTTGACCGGATTAACGAGCAAAATGATAGAGACTTTGAGAGA 1440
1444 GATCACTCCGCTTATGTTGTTGAGACTTGTGCTTACCGCTTGGTGAATTTACGCTGC 1503
1441 GATCACTCCGCTTATGTTGTTGAGACTTGTGCTTACCGCTTGGTGAATTTACGCTGC 1500

QY 1504 GAAAAAGCTAAATGCTTTAAGCGGTGCGGAATATTTTGTGTCCTTTGAGAGAGTGG 1563
DB 1501 GAAAAAGCTAAATGCTTTAAGCGGTGCGGAATATTTTGTGTCCTTTGAGAGAGTGG 1560
QY 1564 AGCATCAGCTCTTGTGGGGGTTTGAACCGGTGAGAGTTATGAAAGAGCGCTTTTAG 1623
DB 1561 AGCATCAGCTCTTGTGGGGGTTTGAACCGGTGAGAGTTATGAAAGAGCGCTTTTAG 1620
QY 1624 AATGACAGCTGCGAGCAAGTATACAGTTTAGATCCTTTTAAATTTCTTTAGCATGA 1683
DB 1621 AATGACAGCTGCGAGCAAGTATACAGTTTAGATCCTTTTAAATTTCTTTAGCATGA 1680
QY 1684 TATTAATCTTAAAGTTTCTCATTTTAAATGATATGTTGTGTGAGAGTTGATCTTTTGTAGCT 1743
DB 1681 TATTAATCTTAAAGTTTCTCATTTTAAATGATATGTTGTGTGAGAGTTGATCTTTTGTAGCT 1740
QY 1744 ACCCCAGCAATATTCACAGAGTATGAAAGTTTGAAGTTGCAATGCTGCTGTGG 1803
DB 1741 ACCCCAGCAATATTCACAGAGTATGAAAGTTTGAAGTTGCAATGCTGCTGTGG 1800
QY 1804 GCTCAAGCTTTTATTTGTAAGAACCAACCTTTTACAGAGTCTGATAGCAATTTCCAG 1863
DB 1801 GCTCAAGCTTTTATTTGTAAGAACCAACCTTTTACAGAGTCTGATAGCAATTTCCAG 1860
QY 1864 CAACCTTCAAGCGCTTAAAGTAAATGCGTATGAGATTCCTGCAATGTTGTATACACGG 1923
DB 1861 CAACCTTCAAGCGCTTAAAGTAAATGCGTATGAGATTCCTGCAATGTTGTATACACGG 1920
QY 1924 AATAATTTGGAGATGACTTGGCTCTAAGAAAGGGGACTGTGCACTGCTTATAGCAAA 1983
DB 1921 AATAATTTGGAGATGACTTGGCTCTAAGAAAGGGGACTGTGCACTGCTTATAGCAAA 1980
QY 1984 GTTGTATGATGCGATATGTTGTGGGCTTACAGAGTGAAGATTCATAATATAGCAATCCA 2043
DB 1981 GTTGTATGATGCGATATGTTGTGGGCTTACAGAGTGAAGATTCATAATATAGCAATCCA 2040
QY 2044 GCTATTTGAGAGTTGTTTGAAGAAATTCAAATCGTATGACAAATGATGATCTCCCTGGA 2103
DB 2041 GCTATTTGAGAGTTGTTTGAAGAAATTCAAATCGTATGACAAATGATGATCTCCCTGGA 2100
QY 2104 CTATGCAATTTGTTGAAAACCTGTGTGCGAGGGGTGCTTCCATAGTTGCAAGAACCC 2163
DB 2101 CTATGCAATTTGTTGAAAACCTGTGTGCGAGGGGTGCTTCCATAGTTGCAAGAACCC 2160
QY 2164 AAAGATTAATAAATTTAACTCGGGGACTACTATGATGATCTATGATTTTGAATCTTG 2223
DB 2161 AAAGATTAATAAATTTAACTCGGGGACTACTATGATGATCTATGATTTTGAATCTTG 2220
QY 2224 GAAAGATGAGAGTATGTTCAAGGTTCTCCTTTAGCTGTGCTGCAACTATGCGAAGAT 2283
DB 2221 GAAAGATGAGAGTATGTTCAAGGTTCTCCTTTAGCTGTGCTGCAAGTATGCGAAGAT 2280
QY 2284 GGAAGCGGAGCATGTGAAAAGTATGCTATGACGAGCACTGCGAAGATTTTCTTCCCGC 2343
DB 2281 GGAAGCGGAGCATGTGAAAAGTATGCTATGACGAGCACTGCGAAGATTTTCTTCCCGC 2340
QY 2344 TATACAGATAGAAACTCGGCTGAACCCAAAGATGTGCAAGAGACAGTGTTTAGTGAAT 2403
DB 2341 TATACAGATAGAAACTCGGCTGAACCCAAAGATGTGCAAGAGACAGTGTTTAGTGAAT 2400
QY 2404 CTTGTTGGTAACTATGATAGAGCCGTATGAGCTGTGCTGCTGTTTATTTGCAAGACTGA 2463
DB 2401 CTTGTTGGTAACTATGATAGAGCCGTATGAGCTGTGCTGCTGTTTATTTGCAAGACTGA 2460
QY 2464 AGACCTCTGAAACCTTTGAAACTATATGATTAATGCAATTTGAGCGGGTCTCAAGAGAT 2523
DB 2461 AGACCTCTGAAACCTTTGAAACTATATGATTAATGCAATTTGAGCGGGTCTCAAGAGAT 2520
QY 2524 AGCGTTGATGAAGAACTACTGTTGAATGTCGTTGCTGATATGTTAAAGAGCAAGTGTG 2583
DB 2521 AGCGTTGATGAAGAACTACTGTTGAATGTCGTTGCTGATATGTTAAAGAGCAAGTGTG 2580
QY 2584 AAGATCTAGCTGCTGTGTGCGAATTGCACTGATTTCACTGTTCAAGCAAGATATTTT 2643

DB 2581 AAGATCTAGCTGCTGTGTGCGAATTGCACTGATTTCACTGTTCAAGCAAGATATTTT 2640
QY 2644 CTTAAGACAGCTCATCTTTTCAACGCAAGGATATAGTTTCTTCTATAGAAATCTGATGC 2703
DB 2641 CTTAAGACAGCTCATCTTTTCAACGCAAGGATATAGTTTCTTCTATAGAAATCTGATGC 2700
QY 2704 GCTACCATAGGATATGATTAATGATGCAATTTTCAATATATCTGCAATGCTCAAAATATGC 2763
DB 2701 GCTACCATAGGATATGATTAATGATGCAATTTTCAATATATCTGCAATGCTCAAAATATGC 2760
QY 2764 TTGTTTGTGAGCTTAAGAAACATATGTTCCACTTAATATACATGCTCCAAAGTTTACCAAG 2823
DB 2761 TTGTTTGTGAGCTTAAGAAACATATGTTCCACTTAATATACATGCTCCAAAGTTTACCAAG 2820
QY 2824 ATTAACAAGTTGCTGAGTAATTTTCACTAATATATGCTGTAATTTTGTATCAAACTG 2883
DB 2821 ATTAACAAGTTGCTGAGTAATTTTCACTAATATATGCTGTAATTTTGTATCAAACTG 2880
QY 2884 TAGACAGAAATGTAATTTTCACTCTCAACATTTCTGTTTAAATATAGATAGATTAAGGA 2943
DB 2881 TAGACAGAAATGTAATTTTCACTCTCAACATTTCTGTTTAAATATAGATAGATTAAGGA 2940
QY 2944 TTGCTTATGTTGCTTGTGCTTGTGCAACTTTTCTTCTGATTTTCTTTTGTGATTTAGG 3003
DB 2941 TTGCTTATGTTGCTTGTGCTTGTGCTTGTGCAACTTTTCTTCTGATTTTCTTTTGTGATTTAGG 3000
QY 3004 GTCAGTCAGAGCTGACATTCAGAAAGCACTCCAGAAATGGAATGCTAGCACTGCAAGAA 3063
DB 3001 GTCAGTCAGAGCTGACATTCAGAAAGCACTCCAGAAATGGAATGCTAGCACTGCAAGAA 3060
QY 3064 TATAGTATCCAAAGTGGCAGAAAGATTAAGTCTCTGGCTTTTGGGCTGATCAACCGCATAGA 3123
DB 3061 TATAGTATCCAAAGTGGCAGAAAGATTAAGTCTCTGGCTTTTGGGCTGATCAACCGCATAGA 3120
QY 3124 AATGTTACCAAGGTGAGGGAATTAATCTACATTTCAATCAATTTGTGAAAACCTGTTGG 3183
DB 3121 AATGTTACCAAGGTGAGGGAATTAATCTACATTTCAATCAATTTGTGAAAACCTGTTGG 3180
QY 3184 ACATGATATATGTCGAGGCTTGTGATTCGTTATTTTATAGTTTGGATGGGCGAA 3243
DB 3181 ACATGATATATGTCGAGGCTTGTGATTCGTTATTTATATATATATATATATATATATATATAT 3240
QY 3244 TGCTGAAGATTTGAGCTGACAGACAGCTGAACTGCGCAGCTTGGGTTGATTAATGAT 3303
DB 3241 TGCTGAAGATTTGAGCTGACAGACAGCTGAACTGCGCAGCTTGGGTTGATTAATGAT 3300
QY 3304 ATPACCTGTTGAATCTATCTGTTGAACAGTGTGACAGTCTGACAGATGAAACCCGTGCTC 3363
DB 3301 ATPACCTGTTGAATCTATCTGTTGAACAGTGTGACAGTCTGACAGATGAAACCCGTGCTC 3360
QY 3364 TGTGTGAAAGCAACTCTGAGAGAGTCTGTTGCTATCTGATTTGATTTGATTCACAAAAA 3423
DB 3361 TGTGTGAAAGCAACTCTGAGAGAGTCTGTTGCTATCTGATTTGATTTGATTCACAAAAA 3420
QY 3424 ATGCTACTGATATGTCAGAACTTACAAACAGATACGAATTTTCTGCTCCAAAGTCAAGGCT 3483
DB 3421 ATGCTACTGATATGTCAGAACTTACAAACAGATACGAATTTTCTGCTCCAAAGTCAAGGCT 3480
QY 3484 GGAATACTGTAAGGCTCTGTTCTTGCAATCAATATATATCTCATATGTAAGATGTCTGAG 3543
DB 3481 GGAATACTGTAAGGCTCTGTTCTTGCAATCAATATATATCTCATATGTAAGATGTCTGAG 3540
QY 3544 CTTGCGAGATTTCTTTGTTCTGTAATTTCTCTCTCAAGTATGATTTTAAATTAAGAA 3603
DB 3541 CTTGCGAGATTTCTTTGTTCTGTAATTTCTCTCTCAAGTATGATTTTAAATTAAGAA 3600
QY 3604 CAAAAAATTAACCTTCTTGCAACACCTTTTCTTGATCTTAACTTAACTATTAAGGCT 3663
DB 3601 CAAAAAATTAACCTTCTTGCAACACCTTTTCTTGATCTTAACTTAACTATTAAGGCT 3660
QY 3664 ACAA 3667
|||||

Db 3661 ACAA 3664

RESULT 3
AB016888
LOCUS
DEFINITION Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:MDH9.
ACCESSION AB016888 BA000015
VERSION AB016888.1 GI:3449329
KEYWORDS
SOURCE
ORGANISM Arabidopsis thaliana (thale cress)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE
AUTHORS 1 Asanizu,E., Sato,S., Kaneko,T., Nakamura,Y., Kotani,H., Miyajima,N. and Tabata,S.
TITLE Structural analysis of Arabidopsis thaliana chromosome 5. VIII. Sequence features of the regions of 1,081,958 bp covered by seventeen physically assigned P1 and TAC clones
JOURNAL DNA Res. 5 (6), 379-391 (1998)
PUBMED 10048488
REFERENCE 2 (bases 1 to 85791)
AUTHORS Nakamura,Y.
TITLE Direct Submission
JOURNAL Submitted (18-AUG-1998) Yasukazu Nakamura, Kazusa DNA Research Institute, Department of Plant Gene Research; 1532-3, Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:ynakamu@kazusa.or.jp, Tel:81-438-52-3935, Fax:81-438-52-3934)
COMMENT Address for correspondence: kaos@kazusa.or.jp
For the latest information on annotation of this clone, please see http://www.kazusa.or.jp/kaos/cgi-bin/agd_graph.cgi?c=MDH9
Genes with similarity to proteins in the databases are described in 'product' or 'note' qualifiers. Genes that have no significant protein similarity are described as 'unknown protein'.
The software programs used to predict genes include: Grail (Informatics Group, Oak Ridge National Laboratory, <http://compbio.ornl.gov/Grail-1.3/>), GENSCAN (Chris Burge, MIT, <http://CCR-081.mit.edu/GENSCAN.html>), NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of Denmark, <http://www.cbs.dtu.dk/services/NetGene2/>) and SplicePredictor (Volker Brendel, Stanford University, <http://grem1ml.zool.iastate.edu/cgi-bin/sp.cgi>).
Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, Washington University School of Medicine, St. Louis, <http://genome.wustl.edu/eddy/tRNAscan-SE/>).
This sequence may not be the entire insert of this clone. It may be shorter because we remove overlaps between neighboring submissions.
The 5' clone is K5J14 and the 3' clone is K16E1.
Location/Qualifiers
1. 85791
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/db_xref="taxon:3702"
/chromosome="5"
/clone="MDH9"
/clone_1fb="Mitcui P1"
/ecotype="Columbia"
/complement(59. 239)
/note="CDS is reported in Acc# AB023032
gene_id:K5J14.12
unknown protein"
/number=1
/evidence=not experimental
complement(join(2056. 2079,2278. 2393,2991. 3633))
/note="contains similarity to unknown protein
gene_id:MDH9.1
pir||T04864"
/pseudo
/codon_start=1
/evidence=not experimental
join(5739. 6138,6462. 6595)

CDS

/note="unnamed protein product; gene_id:MDH9.2
unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAB10474.1"
/db_xref="GI:9759469"
/translation="MKDEPAEEDBKSSNLESLAKNSVPLPERGRFYIEYAKNE
RLRRKGEAEVGGVVGKTPYNLGVNHEPTNRRGTAKESIIRKTVSVSSSIPR
YSLSMAKENRKPPLPLNVAVSAMKSVTYRKEERLLLLCTACNTSSGLAIPRLTGS
EPLGVHKKRTLPSPSALV"
complement(join(6868. 7911,7989. 8414,8555. 8671,
8754. 9137))
/note="gene_id:MDH9.3"
/codon_start=1
/evidence=not experimental
/product="arm repeat containing protein"
/protein_id="BAB10475.1"
/db_xref="GI:9759470"
/translation="WDTEBATGDADENRDEVTAEERIHDEVDAVEIHEEYKEDD
DCBGLVSDIVSIFLDQINGRYRTOQKECFNLVRLKILIPLEDEIRGESPSCKH
LNRLRVKFLAAKLLLETCSNGSKIYALDEETMTFHSIYEKLANVLAPEDEKLI
SGDAKDEIDSLCKOLKAKARTDODIELADVMVVSFSPDNADSAIIERLAKLE
LQTIIDDKETTTAIOSLIODKGLNIEETKHILIELKFKLQLEATDII.YQPVNK
AITSKSLIIPHPFLCPTITELMDPIYITAGQTYKESIQKPFDAHGHTCPKTRQEL
DHLSLAPNFAKLNLMQWCEKNPFKIPKESVDSQSQBQDEVSLVEALSSQLEQ
RSVQKQRLARENENRVLIVANAQIPLVQLSYPDSGIQNAVYTLNLNLSIDEVN
KLLISNGALPNILIEIENGREARENSAALFSLMLDENKVTIGISNGIPLVLDL
OHGELRCKQALFALFSLNSANKGAIIDAGIYQPLNLKDNKLGMIIDEALSLIL
LASHPEBROALIGLSPLETVEFTIROTGPKNKCAVSLLEGSNNSSFIILALQGV
YEVLETTSCTYRQAKNALIQLSKSQI"
10888. 12579
/note="unnamed protein product; gb|AAF07802.1
gene_id:MDH9.4
similar to unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAB10476.1"
/db_xref="GI:9759471"
/translation="MMISEKPLGEBESIRQDLEVLTVSRRLVKSVSQKLKKIKHTEV
VEBEIRGAVNCLISVGCRAVDGTGDEPDSNKKWSASBEGKGLMTICGTEETL
DCFSYGRERFMKNNRKYLADSGODYRKHVYLPDILMCLMRPLTSLNLNHLVCK
KMSMANOTFLMRREGSFOTPMFLPALKXGSCSGSDIHYGVDSQDKRI.ETDL
KGRPMYSVTSIHIEIYIVGSRSDRNSFKSHRGLIYFSPSIIKAKRKIASRRARS.LPI
VGATEVSEPTVQTKONRODRPHLSRVGSESDVDEPRLSVRONRNSADQNGTK
SHRLIRKQRLNRNNSKRFVLIAIGTGI.FDEPLDSGEIYDSATNTWSBMO.LPMGF
GVVSCGIIICNGIFAYASENDKLSGYIDERGFWITIGTSP.PPVNHEFYPLVACNHL
FMLSVCSDGSDGIGBRNKAVKRLMLDLVLTWTEVSVHPDAPMDMNTATYSDQVI
LMGEMFKIRGQVLSFPTVCDVLTESAMHVSBRNQSKLNTSCNKTIALHLH"
complement(15087. 16778)
/note="unnamed protein product; gb|AAF07802.1
gene_id:MDH9.5
similar to unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAB10477.1"
/db_xref="GI:9759472"
/translation="MMISEKPLGEBESIRQDLEVLTVSRRLVKSVSQKLKKIKHTEV
EDEEVARGAVNCLISVGCRAVDGTGDEPDSNKKWSASBEGKGLMTICGTEETL
DCFSYGRERFMKNNRKYLADSGODYRKHVYLPDILMCLMRPLTSLNLNHLVCK
KMSMANOTFLMRREGSFOTPMFLPALKXGSCSGSDIHYGVDSQDKRI.ETDL
KGRPMYSVTSIHIEIYIVGSRSDRNSFKSHRGLIYFSPSIIKAKRKIASRRARS.LPI
VGATEVSEPTVQTKONRODRPHLSRVGSESDVDEPRLSVRONRNSADQNGTK
SHRLIRKQRLNRNNSKRFVLIAIGTGI.FDEPLDSGEIYDSATNTWSBMO.LPMGF
GVVSCGIIICNGIFAYASENDKLSGYIDERGFWITIGTSP.PPVNHEFYPLVACNHL
FMLSVCSDGSDGIGBRNKAVKRLMLDLVLTWTEVSVHPDAPMDMNTATYSDQVI
LMGEMFKIRGQVLSFPTVCDVLTESAMHVSBRNQSKLNTSCNKTIALHLH"
complement(join(18111. 18193,18281. 18535,18773. 19001,
19088. 19195,19307. 19447,19542. 19952,20129. 20245))
/note="unnamed protein product; gene_id:MDH9.6
similar to unknown protein
sp|P42251"
/codon_start=1
/evidence=not experimental

/protein_id="BAB10478.1"
/db_xref="GI:9759473"
/translation="MNQISFMCISLLITVIAAEETPTIRIAFGSCANQSAPOPIWDA
INKPBPOLFIWLDNIGYDRIIRPLVIIGERTFGPRNRPPLVPSSEEMKLRVAKK
ANPXSRIORAKKAGIGTMDHDYGLNDACKERDRKYNOKMLDPIEPLDSDPRKAK
GVAAYTVGPNRKYKVVITDTRHYRDLRSDGSLIGDPQWMLSELSRPSSETII
GSSVQVYISLSTTGPFLPYMESWGRPPKRRKFLKILADTKDQVITFISGDVHSEIT
RYDCSVGPLYDVYSSGLVQSVKVPRLRIIVRLFEYTSITKRVINDNCKFKSCT
YQONFGAISIDMNAIPVITKLEIRVDNDTVIGTVNLSSEIQPGGNSLKTITTKGK
SQRCTLEIELEPPTRRYLAIVLYFTIATLMAIIGLIGAVIATACVYCKKVD"
complement(21540..22097)
/note="unnamed protein product; contains similarity to
calmodulin
gene_id:MDH9.7"
/codon_start=1
/evidence=not experimental
/protein_id="BAB10479.1"
/db_xref="GI:9759474"
/translation="MTLAKQKSSLSRLYKYSKRSESRNLEDESRSSSSSS
LAVNELRTVFDVMDANSDKGISGEELQCVSLGGALSGREVEVVKTSVDVGDGFI
FEEFLKMEGSDSEERKELKEAFGMVMEGEERTTAASIRRTLRLSGESCIVDAC
KMIKGFNDGDVLSDFEFLMMR"
join(24090..24289,24410..24614,24701..24945,25115..25313,
25632..25679,25821..25928,26039..26227,26314..26613,
26730..26831,27169..27303,27487..27711,27919..28122,
28213..28380,28646..28795,28889..29026,29204..29432,
29519..29671,29784..29910,30115..30241,30554..30674,
30763..30843,30932..31033,31112..31262,31338..31429)
/note="gene_id:MDH9.8"
/codon_start=1
/evidence=not experimental
/product="pili1lysin"
/protein_id="BAB10480.1"
/db_xref="GI:9759475"
/translation="MASSSSSIFGVKSPILAPNSGDSRPRYLKDSRNKVRNPS
SRRLTPHRVYAPSLIIPYNGIMAAOPNSHKRLKRNYSGEATGISLSQRNCFIT
CKRNOAGIRALPSAFVDRTAFLSRSSITSLRHSQIVNATLIDPDEPHAQVAMP
GIVARODLIDLEPDISALEBAFLGCELPSPKHLRQGLGKNLKYLLIPNVPNRF
BAHMEVHVSIDIEBDEQGIAMIEHVAFLGSKREKILGTGRANAAVITDFHVTPI
HSPTKDSBEDLPSVLDALNEIAHPFLSRRVEKERBALISELMMNTTSEIVYDC
OLLOLHSENKIGRRPPIGLEROIKKMDVDRKPKHERMYPPANATLYIVGDINDIR
LYHNIENYGRKGLDNESSTPSSPSPGAPMANFLVPLKIPAGLGTFPSNEKTINDOS
KMITKBRHAIIRPVENHNSLPGTSVLDLPPOIFKHILQNFALNFCPIPVKVTGF
DLRNLMKRIFLSAHFIRINTRYKSNDEPFTVELDHDSGSEGTVTLLITVAPON
MONAVYVAQVEARLKEFGVTGELTRVMDALKDSSEHIAAINDVSVNDNDFMES
DALSHVMDQTOGHETLVAVAGTVLRENTVGAULVEITISPGTAPLPAIYACV
PKYHVDGSESDPNISPDIEISVSKGILAEIPREIYEKELIOSOLKEILFOR
NCPYPIRSGSITKXHDKEGTITQRLSGLIAVNYKKTSTESRAEMRLIVGGAAE
TSDSGAVVVGITLSBGRVDFSRBOVELFCVNLINCSLESTEEFLAMEFRFLR
DNGMOAARQLMLVLERISVLEDAFDRAQLYSYRSIPKLEBETAKMLAMING
DERFVEPTPKSLINLESVDVAMSHFVGNMEVSIQVDFSEBETIRCIIDLYGTVK
ASHDSAKPPGSEPILEPROPTAGLOOQVLEKOTDRACAYIINGPMPNMGVDFDIL
FOSVSKLPAHNGILKSEBOLLEGGDRLOKRAHPLFCVGTMLAEITINSRPFT
VDSIGLITVDSFEINLPRLNIGVIVISVSTPKGVKANDACSVLRGLHSNDIAP
RELDAKRTLLMRHAEIKSNAYWNLNLAHQASSVPKRELSICIVLSLEAASIED
IYLANOLRVDEDSLYSCIGIAGAQAAGEITVLSEEBEBEDVPSCEIVSVGRSSMTTR
PTT"
complement(join(33251..33326,33512..33663,33744..33977,
33888..33968,34143..34265,34359..34768,34873..35052,
CDS
Query Match 99.8%; Score 3659; DB 15; Length 85791;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 3662; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 TGTTCGATTAAGAGATTAATTAAGCAATTTGCTTGATTTCAACAAGATTTTG 60
DB TGTTCGATTAAGAGATTAATTAAGCAATTTGCTTGATTTCAACAAGATTTTG 63656
QY 61 CTTGGCTAAGATTCATGCTCTGTTGCTTTTACATTTACATGCTCATATAGTTTCG 120
DB CTTGGCTAAGATTCATGCTCTGTTGCTTTTACATTTACATGCTCATATAGTTTCG 63716
QY 121 AATTTTACATTTTCACTTGATGTTTAAAGAAAAGAGGAAATTGATGGGGTTTGTGGG 180

DB 63717 AATTTTACATTTTCACTTGATGTTTAAAGAAAAGAGGAAATTGATGGGGTTTGTGGG 63776
QY 181 TTTAACTTTAAAGTAGTCAAGAAATTAGTCAATGTTTACTGTTGCTCTATATGTTAA 240
DB TTTAACTTTAAAGTAGTCAAGAAATTAGTCAATGTTTACTGTTGCTCTATATGTTAA 63836
QY 241 AATGAAGCAACTCCAACGGTTCTTAGGTGAATATGATATTTAGACGATTTAACATCAT 300
DB AATGAAGCAACTCCAACGGTTCTTAGGTGAATATGATATTTAGACGATTTAACATCAT 63896
QY 301 AAAGTCCTGGCGACTGTAAACATCATATGTTTATTTTATTTTTCAGTAGCTGTGAT 360
DB AAAGTCCTGGCGACTGTAAACATCATATGTTTATTTTATTTTTCAGTAGCTGTGAT 63956
QY 361 GTTTTGTATTTAACTTTATCTACTCAAAATCAAAATTCATTAACCTTAGACGACCAA 420
DB GTTTTGTATTTAACTTTATCTACTCAAAATCAAAATTCATTAACCTTAGACGACCAA 64016
QY 63957 GTTTTGTATTTAACTTTATCTACTCAAAATCAAAATTCATTAACCTTAGACGACCAA 64016
DB 64017 CAGTCTCTTCAATATATGTAAGAAAGAAAGTTTGTAGTAGCCTTAAAGAACACTCC 64076
QY 481 ATGAAGCTCTGAGTCACTGCGGCACTTGGTCTCTCCCATTCGATTTATGCCGATTACA 540
DB ATGAAGCTCTGAGTCACTGCGGCACTTGGTCTCTCCCATTCGATTTATGCCGATTACA 64136
QY 64077 ATGAAGCTCTGAGTCACTGCGGCACTTGGTCTCTCCCATTCGATTTATGCCGATTACA 64136
DB 541 CCGGAGACGACAAAGCTCGGACGTAGGACCAACACCTTACAACTATGTGCTCGCGCAGC 600
QY 64137 CCGGAGACGACAAAGCTCGGACGTAGGACCAACACCTTACAACTATGTGCTCGCGCAGC 64196
DB 601 AAATGGCCGACCGCTCTCTCCGACTTCAATTTCACTCGGATTCCTCTCTCTCTCC 660
QY 64197 AAATGGCCGACCGCTCTCTCCGACTTCAATTTCACTCGGATTCCTCTCTCTCTCC 64256
DB 661 TTCCGACCGGACCAACCAACCGGCACTCTGCTCTCTGCGCACTTATGATGTCTCC 720
QY 64257 TTCCGACCGGACCAACCAACCGGCACTCTGCTCTCTGCGCACTTATGATGTCTCC 64316
DB 721 GAAGCGCAGTCCCATCCCATGATTTCTTACACAGGATTAAGGACTCAAAACATTTTC 780
QY 64317 GAAGCGCAGTCCCATCCCATGATTTCTTACACAGGATTAAGGACTCAAAACATTTTC 64376
DB 781 TTAACCGATGATCAGAAAGAGCAATTCAGACGAGGTTTGGAAACCGCGCAATTCGCT 840
QY 64377 TTAACCGATGATCAGAAAGAGCAATTCAGACGAGGTTTGGAAACCGCGCAATTCGCT 64436
DB 841 TTACCGACGACGCTTTATATGACCGGAGACAGATTTTCAAGCTGCTTGGAAACTCTG 900
QY 64437 TTACCGACGACGCTTTATATGACCGGAGACAGATTTTCAAGCTGCTTGGAAACTCTG 64496
DB 901 TCTATCTCGGCTAGAAAGAGAGTCAATGAAGGCTCTTCTGATGAGAAAGAGCTACA 960
QY 64497 TCTATCTCGGCTAGAAAGAGAGTCAATGAAGGCTCTTCTGATGAGAAAGAGCTACA 64556
DB 961 GTCATCACTGATGTTCTTGGGATTAAGGTAATTTGATTCGAAATATAAGTTTCTTC 1020
QY 64557 GTCATCACTGATGTTCTTGGGATTAAGGTAATTTGATTCGAAATATAAGTTTCTTC 64616
DB 1021 GTTTTAAATTTCAATGATTTGATTAAGAAAGAACTTTTATCTAGTGAAGTTCTGGGCG 1080
QY 64617 GTTTTAAATTTCAATGATTTGATTAAGAAAGAACTTTTATCTAGTGAAGTTCTGGGCG 64676
DB 1081 TCTCGGTATTTGCAAGAAAGGTGTGAGCTGAGATGTTCTTGGGCTGGTGTGAGGCTCT 1140
QY 64677 TCTCGGTATTTGCAAGAAAGGTGTGAGCTGAGATGTTCTTGGGCTGGTGTGAGGCTCT 64736
DB 1141 GCTTAAGAGAGGTGCTTAAGTCTTAAGCAAGATGTGTTTATGATTAAGGCGCTTGC 1200
QY 64737 GCTTAAGAGAGGTGCTTAAGTCTTAAGCAAGATGTGTTTATGATTAAGGCGCTTGC 64796
DB 1201 GTTCTGATGTCTCGAGGATGCTATGAGCAATTTGATTCACCTGATTTATTAATCGGTTA 1260

Dh 64797 GTTTCGATGCTCGAGAGATGCTATGCAATTGATCCACCTGATTTTATTAACGTGTTA 64856
Qy 1261 TGAAGTTGTTGAGAAAGCTTTTGAAGCTTTTACAGGTAGTTGACTTGTGCTTGGTAATTG 1320
Db 64857 TGAAGTTGTTGAGAAAGCTTTGAAAGCTTTTACAGGTAGTTGACTTGTGCTTGGTAATTG 64916
Qy 1321 ACAGACGTTGGCTTTTATTAAGAACTTTCTGTGATTTTGATACTTTGTTATTTAGTCTTGTGTA 1380
Db 64917 ACAGACGTTGGCTTTTATTAAGAACTTTCTGTGATTTGATACTTTGTTATTTAGAGCTTGTGTA 64976
Qy 1381 GGAGGAGAGAGCAAGTACCTTGTGACCGGATTTAGTGCACAAATGATGAGACTTTGGA 1440
Db 64977 GGAGGAGAGAGCAAGTACCTTGTGACCGGATTTAGTGCACAAATGATGAGACTTTGGA 65036
Qy 1441 AGAGATCACTCCGCTTATGCTTGTGAGACTTGTGGCTTACCGCTGGTGTGATGATTAAGC 1500
Db 65037 AGAGATCACTCCGCTTATGCTTGTGAGACTTGTGGCTTACCGCTGGTGTGATGATTAAGC 65096
Qy 1501 TGGCAAAAGACTAAATGCTTTAAGCGGTGTGCGAAATATTTTGTGCTGTGAGAGAG 1560
Db 65097 TGGCAAAAGACTAAATGCTTTAAGCGGTGTGCGAAATATTTTGTGCTGTGAGAGAG 65156
Qy 1561 TGGAGCATCAGCTCTTGTGTGGGGTTTGAACCGGTGAGAAAGTTTATGATGAGAGCGTTT 1620
Db 65157 TGGAGCATCAGCTCTTGTGTGGGGTTTGAACCGGTGAGAAAGTTTATGAAATGAGCGTTT 65216
Qy 1621 ACGAATGACAGCTGTGAGAGAGTATACAGTTTATACCTTTTATTTTATTTCTTTAGCA 1680
Db 65217 ACGAATGACAGCTGTGAGAGAGTATACAGTTTATACCTTTTATTTTATTTCTTTAGCA 65276
Qy 1681 TGAATTAACCTTAAAGTTTCTCATTTTAAATGTAATGTTGTGTGAGTTGATCTTTTGTGA 1740
Db 65277 TGAATTAACCTTAAAGTTTCTCATTTTAAATGTAATGTTGTGTGAGTTGATCTTTTGTGA 65336
Qy 1741 GCTACCCCAAGAAATNTCCAGAGAGTCAATTGATGAAAGTTTACGAAGTTGCCTTGTCT 1800
Db 65337 GCTACCCCAAGAAATNTCCAGAGAGTCAATTGAAAGTTTACGAAGTTGCCTTGTCT 65396
Qy 1801 GTGGCTCAGCTTTTATTTGTTAGTAAGCAACCTTTTACAGSAGTGTGATAGCAATT 1860
Db 65397 GTGGCTCAGCTTTTATTTGTTAGTAAGCAACCTTTTACAGSAGTGTGATAGCAATT 65456
Qy 1861 CAGCAACTTCAGCAGGCTTAAGGTAAATGCTATGAGATTCCTGCGATGTTGATGATACA 1920
Db 65457 CAGCAACTTCAGCAGGCTTAAGGTAAATGCTATGAGATTCCTGCGATGTTGATGATACA 65516
Qy 1921 CGGAATTAATTTGGAGATAGACTTGGCTCTAGAAAAGGGACCTGTGTGCACTGCTTATAGG 1980
Db 65517 CGGAATTAATTTGGAGATAGACTTGGCTCTAGAAAAGGGACCTGTGTGCACTGCTTATAGG 65576
Qy 1981 AAAGTTGATGAAATGCGTATGTGTGGGCTTGAACAGTGAAGSATTCAATATATAGAA 2040
Db 65577 AAAGTTGATGAAATGCGTATGTGTGGGCTTGAACAGTGAAGSATTCAATATATAGAA 65636
Qy 2041 CCAAGCTATTTGGAGTTTGTGTTGAGAAATTCAAATCGTATGACAAATGATGATCTCCT 2100
Db 65637 CCAAGCTATTTGGAGTTTGTGTTGAGAAATTCAAATCGTATGACAAATGATGATCTCCT 65696
Qy 2101 GGACATGACAAATTTGTTGAGAAACCTGTGTGCAAGGGGTTGTCTTCCAGAGTTCAAGAC 2160
Db 65697 GGACATGACAAATTTGTTGAGAAACCTGTGTGCAAGGGGTTGTCTTCCAGAGTTCAAGAC 65756
Qy 2161 ACCAAAGATTAATAAATTTTAACTCGGGGACTACTATGATGATCCATGTTTGTAGTAC 2220
Db 65757 ACCAAAGATTAATAAATTTTAACTCGGGGACTACTATGATGATCCATGTTTGTAGTAC 65816
Qy 2221 TTGGAAGAGTGAAGTAGTTTACAGGTTCTCTTTTAAAGTGTCTGTCAACATATGCAAGG 2280
Db 65817 TTGGAAGAGTGAAGTAGTTTACAGGTTCTCTTTTAAAGTGTCTGTCAACATATGCAAGG 65876
Qy 2281 ATTGGAAGCCGAGCATGTGAAAAGTATGCTATGACGAGCATGCAAGAAAGTTTCTCTCC 2340
Db 65877 ATTGGAAGCCGAGCATGTGAAAAGTATGCTATGACGAGCATGCAAGAAAGTTTCTCTCC 65936

Qy 2341 CGCTATACAGATAGAAACTCGGCTGAACCCAGAGATGTGCAGAGACAGTGTATGATGA 2400
Db 65937 CGCTATACAGATAGAAACTCGGCTGAACCCAGAGATGTGCAGAGACAGTGTATGATGA 65996
Qy 2401 GATCTGTGTGTAACAATGTAGAGCGGTGATGTGAGGCTGTGTCTTTTATTTGCAAGAGCT 2460
Db 65997 GATCTGTGTGTAACAATGTAGAGCGGTGATGTGAGGCTGTGTCTTTTATTTGCAAGAGCT 66056
Qy 2461 GTAAAGCCCTGTGAAAACCTTGAACCTAATGATTAATGCAATTTGAGGCTGAGCTCAGAG 2520
Db 66057 GTAAAGCCCTGTGAAAACCTTGAACCTAATGATTAATGCAATTTGAGGCTGAGCTCAGAG 66116
Qy 2521 AGTAGCCTTATGATGAACCTACTGTGTAATGTCCGTGCTGATATGTTTAAAGAGGCAAGT 2580
Db 66117 AGTAGCCTTATGATGAACCTACTGTGTAATGTCCGTGCTGATATGTTTAAAGAGGCAAGT 66176
Qy 2581 GTGAAGTCCTAGCTGTGTGTGTGCAATTGAGCTGATTTCACTGTTCAAGCAGAGAT 2640
Db 66177 GTGAAGTCCTAGCTGTGTGTGTGCAATTGAGCTGATTTCACTGTTCAAGCAGAGAT 66236
Qy 2641 TTTCTTAAAGCAGCTCATCTTTTCAAGCAGAGATATGTTTCTTCTATGAAATCTGAT 2700
Db 66237 TTTCTTAAAGCAGCTCATCTTTTCAAGCAGAGATATGTTTCTTCTATGAAATCTGAT 66296
Qy 2701 GTCGCTACCAATAGTATGATTAATGATGCAATTTTCAATATCTGATTCCTCAAAATA 2760
Db 66297 GTCGCTACCAATAGTATGATTAATGATGCAATTTTCAATATCTGATTCCTCAAAATA 66356
Qy 2761 TGTCTGTTTGTGTAGCTTAAGAAATAGTTCCTCAATATCANTGCCCAAAAGTTGATCC 2820
Db 66357 TGTCTGTTTGTGTAGCTTAAGAAATAGTTCCTCAATATCANTGCCCAAAAGTTGATCC 66416
Qy 2821 AAGATTAACAAGTTGCTGAGTAATTTCACTAATATAGCTGTGAATTTTGTGATCAAA 2880
Db 66417 AAGATTAACAAGTTGCTGAGTAATTTCACTAATATAGCTGTGAATTTTGTGATCAAA 66476
Qy 2881 CTGTAGACAGAAATGTAAATTTTCACTCTCAAACTTCTGTTTGAATTAAGTATGATAG 2940
Db 66477 CTGTAGACAGAAATGTAAATTTTCACTCTCAAACTTCTGTTTGAATTAAGTATGATAG 66536
Qy 2941 AGATTGCTTATGTGTGCTTGTGTCCAACTTTTCTTCTGATTTTCTTTTCCATTT 3000
Db 66537 AGATTGCTTATGTGTGCTTGTGTGTCCAACTTTTCTTCTGATTTTCTTTTCCATTT 66596
Qy 3001 AGGCTCAGCAGAGCTGACGATTCAGAGACCTTCCAGAAAGGATGCTAGAGACTGACAG 3060
Db 66597 AGGCTCAGCAGAGCTGACGATTCAGAGACCTTCCAGAAAGGATGCTAGAGACTGACAG 66656
Qy 3061 GAATATAGTATCCAAAGTGTGCAAGAAATTAAGTCTGTGGCTTTTGGGCTGTATCACCGCAT 3120
Db 66657 GAATATAGTATCCAAAGTGTGCAAGAAATTAAGTCTGTGGCTTTTGGGCTGTATCACCGCAT 66716
Qy 3121 AGAAATGTTACAGAGGTGAGGGAAATTAATCTTCAATTCATTAATTTGTGTGAAAATCTGT 3180
Db 66717 AGAAATGTTACAGAGGTGAGGGAAATTAATCTTCAATTCATTAATTTGTGTGAAAATCTGT 66776
Qy 3181 TGAACATGATTAATAGTGTGTGCTGTTGATTTCTGTATTAATAGTTTGTGATGGG 3240
Db 66777 TGAACATGATTAATAGTGTGTGCTGTTGATTTCTGTATTAATAGTTTGTGATGGG 66836
Qy 3241 GAATGCTGAAGATTTTGAACCTGACAGAGCAGCTGAACCTGCGAGCTTGGGTTGTTATG 3300
Db 66837 GAATGCTGAAGATTTTGAACCTGACAGAGCAGCTGAACCTGCGAGCTTGGGTTGTTATG 66896
Qy 3301 ATTATACACTGTGTAACCTATCTGTTGACAGTGTGACAGTCTCAGCAGATGAGAACCCGTG 3360
Db 66897 ATTATACACTGTGTAACCTATCTGTTGACAGTGTGACAGTCTCAGCAGATGAGAACCCGTG 66956
Qy 3361 CTGTGTGGAAGCAACTGTGGAAGAGTCTGCTTGTCTATCTGATTTGGTTCAATCCAGAAA 3420
Db 66957 CTGTGTGGAAGCAACTGTGGAAGAGTCTGCTTGTCTATCTGATTTGGTTCAATCCAGAAA 67016

QY 3421 ACAATGCTAGTATGTCAGAACCTTACACAAAGATGCAAGTTTCTGCTCCAGTONG 3480
DB 67017 ACAATGCTAGTATGTCAGAACCTTACACAAAGATGCAAGTTTCTGCTCCAGTONG 67076
QY 3481 GGTGGAATCACTGAGGCTGCTTCTGATCAATATATCTCATATGATGCAATGCT 3540
DB 67077 GGTGGAATCACTGAGGCTGCTTCTGATCAATATATCTCATATGATGCAATGCT 67136
QY 3541 GAGCTTGCAGATTCCTCTTCTGTAATTAATCTCTCTTAAGTTAGTTTATAATGA 3600
DB 67137 GAGCTTGCAGATTCCTCTTCTGTAATTAATCTCTCTTAAGTTAGTTTATAATGA 67196
QY 3601 ACACAAAAAATTAAGCTTCTTGGACACACCTTTCTTGTATCTAACTATAACATAG 3660
DB 67197 ACACAAAAAATTAAGCTTCTTGGACACACCTTTCTTGTATCTAACTATAACATAG 67256
QY 3661 GCTACAA 3667
DB 67257 GCTACAA 67263

RESULT 4
AY091075
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
COMMENT

AY091075 2637 bp mRNA linear PLN 18-SEP-2002
Arabidopsis thaliana unknown protein (At5g42480) mRNA, complete cds.
AY091075.1 GI:20259550
FLI CDNA.
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Buxaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eustrods II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 2637)
Yamada, K., Banh, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Liu, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shim, P., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.
Arabidopsis Full Length cDNA Clones
Unpublished
2 (bases 1 to 2637)
Yamada, K., Banh, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Liu, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shim, P., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.
Direct Submission
Submitted (21-MAR-2002) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA
RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN Arabidopsis Full-length cDNA': Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.
The Sak, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Yamada, K., Banh, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B., Liu, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shim, P., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.
Yamada, K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally to

this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGEC) contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis genome submitted to GenBank.

FEATURES

source

1. 2637

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/chromosome="5"

/clone="RAFL09-76-G11 (R19395)"

/ecotype="Columbia"

/note="This clone is in a modified pBluescript vector (pLC-1) as a BamHI/XhoI insert."

1. 2637

/gene="At5g42480"

1. 113

/gene="At5g42480"

114. 2519

/gene="At5g42480"

/codon_start=1

/evidence="experimental"

/product="unknown protein"

/protein_id="AA013895.1"

/db_xref="GI:20259551"

/translation="MEALSHVIGIGLSPOLCRPLPATTKLRSHNTSTTICSAKMDRLSDPNTSDSSSSPATATTTATVATVPPSISREHNVIPIDFVGLAQHFLTDTIRAPRARSPPQRFSDALISRQIIQAACRTISRSRNRVNEGILDDSEATVITVPMDKVPALCVLQEGEETELVGLALKERIPKSPKQDVVLMALVGLDVSRLDAAADPPDFITGEFVEALKLOEGEASLAPDLAQIDLEETTPRAVILGLPDLDDYAKRLNGISGVRLNLSVGGASALVGLTRERFMRNAPFRMTABQVLDVATPNSNPASSPEVEVALVVAQAFIGKKPHLLDADKFOOLQOAKVAMEIPAMLYDRNMWEIDFGLEIGICALLIGKVECMWILGIDSDSQRNPALIVEPVENSRDNDLRLGCKLLETMTAGVYPPRRDTODKPKGDIYDDPWVLSYLERVVOGSPPLAALAMRITGAEHYKASQMLQVFFSRITDRKSAEPDVQETVFSVDPVGNNGRDGEPEVFLAEAVRPSNFEPTNDYALRAVSESSVDETVMESVADMLKEASVYLIAAGVIGISLSPQSYQFLPKSSSPQRKDMVSGESDVAITGSRVADSDSEALPRMARTENIUSKQKIKSLAFGPDHRIEMLPVLDGMLKITDRAETQGLGLVDTTLKLSVDSVTVASDGRALVEATLSESAQLSLVHENNATVRYTTRTYEVWMSGKMTGSGVLAS"

2520. 2637

/gene="At5g42480"

ORIGIN

3'UTR

Query Match 52.7%; Score 1931; DB 15; Length 2637;
Beet Local Similarity 80.4%; Pred. No. 0;
Matches 2616; Conservative 0; Mismatches 5; Indels 632; Gaps 5;

QY 368 GATTAACTTATCTACTCAAAATCAAAATCCATAAACCCTAGACGACAAACAGTCTC 427
DB 1 GATTAACTTATCTACTCAAAATCAAAATCCATAAACCCTAGACGACAAACAGTCTC 60
QY 428 TTCAATATGTAAACAGAACAAAGTTTGTAGTAGCTTAAAGACACTCCCATGAG 487
DB 61 TTCAATATGTAAACAGAACAAAGTTTGTAGTAGCTTAAAGACACTCCCATGAG 120
QY 488 CTCTGAGCAGTGGGAGCTTGTCTCCCATTCATTTAGTTCATTCACCGGGA 547
DB 121 CTCTGAGCAGTGGGAGCTTGTCTCTCCCATTCATTTAGTTCATTCACCGGGA 180
QY 548 CGAACAAGCTCGAGCTGACCAACACCTTACAACTATGCTCCGCCAGCAATGG 607
DB 181 CGAACAAGCTCGAGCTGACCAACACCTTACAACTATGCTCCGCCAGCAATGG 240
QY 608 CCGACCGTCTTCTCTCGCATTTCAATTTACCTCGATTCCTCTCTCTCTTGGCCA 667
DB 241 CCGACCGTCTTCTCTCGCATTTCAATTTACCTCGATTCCTCTCTCTCTTGGCCA 300
QY 668 CCGCACACACACCGGCACTGCTCTGCGACCATTAATGATGTCGCGAAGGCC 727
DB 301 CCGCACACACACCGGCACTGCTCTGCGACCATTAATGATGTCGCGAAGGCC 360

QY 728 ACCTCCCATCCCATTTGATTTCTACAGAGTATTAAGAGCTCAACACATTTCTTAACCG 787
Db 361 ACGTCCCATCCCATTTGATTTCTACAGAGTATTAAGAGCTCAACACATTTCTTAACCG 420
QY 788 ATGAATAGAAAGACATTCGAAAGCTAGAGTTTGAAAACCGCCGAATTCGGTTTCAGCG 847
Db 421 ATGAATAGAAAGACATTCGAAAGCTAGAGTTTGAAAACCGCCGAATTCGGTTTCAGCG 480
QY 848 ACGAGCTTTATACAGCCGAGACAGATTTCTTCAAGCTCTTGCGAAACTCTGTCTATC 907
Db 481 ACGAGCTTTATACAGCCGAGACAGATTTCTTCAAGCTCTTGCGAAACTCTGTCTATC 540
QY 908 CTCGGTCTAGAAGAGATTAATGAAGGCTCTTCTGATGATGAAGAGCTACAGTCAATCA 967
Db 541 CTCGGTCTAGAAGAGATTAATGAAGGCTCTTCTGATGATGAAGAGAGCTACAGTCAATCA 600
QY 968 CTGATGCTCTTGAGATTAAGGTAATTTGATTTGCGAATTAATAAGTTTCTTCGTTTAA 1027
Db 601 CTGATGCTCTTGAGATTTGCGAATTTGCGAATTAATAAGTTTCTTCGTTTAA 617
QY 1028 TTTCAATGAATTTGATAAAGAAAGAACTTTTATCTAGTGAAGGTTCTGAGGCTCTCTGT 1087
Db 618 -----AAGTTCTCTGTCTCTGT 638
QY 1088 GTATTTGCAAGAAGGCTGTGAGACTGAGATAGTTCTTCGCGTTGTGAGGCTCTGCTTAAG 1147
Db 639 GTATTTGCAAGAAGGCTGTGAGACTGAGATAGTTCTTCGCGTTGTGAGGCTCTGCTTAAG 698
QY 1148 GAGAGGTTGCTTAAGTGTGTTTAAGCAAGATGCGTTTATGTTAATGAGCGCTTCGTTCTC 1207
Db 699 GAGAGGTTGCTTAAGTGTGTTTAAGCAAGATGCGTTTATGTTAATGAGCGCTTCGTTCTC 758
QY 1208 GATGCTTCGAGGAGATGCTATGAGCATTTGATCCACCTGATTTTATCTGTTATGAGTTT 1267
Db 759 GATGCTTCGAGGAGATGCTATGAGCATTTGATCCACCTGATTTTATCTGTTATGAGTTT 818
QY 1268 GTTGAAGAAGCTTTGAAGCTTTTACAGGTAAGTTGACTTCTTGTAATTTGACGACCG 1327
Db 819 GTTGAAGAAGCTTTGAAGCTTTTAC----- 843
QY 1328 TTGGCTTATAAGAACTTTCTTGAATTTGATCTTTGTAATGAGTCTTGTAAGGAGAA 1387
Db 844 -----AAGAGAA 851
QY 1388 GGAGCAAGTACCTTTGCAACCGGATTTACGTGCACAAATTGATGAGACTTTGGAAGAGATC 1447
Db 852 GGAGCAAGTACCTTTGCAACCGGATTTACGTGCACAAATTGATGAGACTTTGGAAGAGATC 911
QY 1448 ACTCCGCTTATGTTTGAAGCTACTTGGCTTACCGCTTGATGATTAAGCTGCGAAA 1507
Db 912 ACTCCGCTTATGTTTGAAGCTACTTGGCTTACCGCTTGATGATTAAGCTGCGAAA 971
QY 1508 AGACTAAATGTTTAAGCGGTGTCGGAATTTTGTGTCGTGTTGAGAGAGGTGAGACA 1567
Db 972 AGACTAAATGTTTAAGCGGTGTCGGAATTTTGTGTCGTGTTGAGAGAGGTGAGACA 1031
QY 1568 TCAGCTCTTGTGGGGGTTTGACCCGTGAGAAATTATGAATGAGGCTTTTTCAGATG 1627
Db 1032 TCAGCTCTTGTGGGGGTTTGACCCGTGAGAAATTATGAATGAGGCTTTTTCAGATG 1091
QY 1628 ACAAGCTGCTAGCAGATTAACAGTTTAGATACCTTTTAAATTTCTTTAGCAGATATTA 1687
Db 1092 ACAAGCTGCTAGC----- 1104
QY 1688 ACTTAGGTTTCTCATTTTAATGTAATGTTGTGTGTAAGTTGATCTTTTGTAGTACC 1747
Db 1105 -----AGTTGATCTTTTGTAGTACC 1128
QY 1748 CAAGCAATATTCAGCAGAGTCAATTTGAAGTTTACGAAGTTGCACTTGTCTGTTGAGCTC 1807
Db 1129 CAAGCAATATTCAGCAGAGTCAATTTGAAGTTTACGAAGTTGCACTTGTCTGTTGAGCTC 1168
QY 1808 AAGCTTTTATTTGTAAGAACCAACCTTTTACAGATGCTGATTAAGCAATTCAGCAAC 1867

Db 1189 AAGCTTTTATTTGTAAGAACCAACCTTTTACAGATGCTGATTAAGCAATTCAGCAAC 1248
QY 1868 TTCAGAGGCTAAGGTAATGAGCTATGAGAGATTCCTGCGATGTTGTATGATACAGGAATA 1927
Db 1249 TTCAGAGGCTAAGGTAATGAGCTATGAGAGATTCCTGCGATGTTGTATGATACAGGAATA 1308
QY 1928 ATTGGAGATTAAGCTTCGCTCTAGAAAAGGAGACTCTGTGCACTGCTTATAGGCAAGTTG 1987
Db 1309 ATTGGAGATTAAGCTTCGCTCTAGAAAAGGAGACTCTGTGCACTGCTTATAGGCAAGTTG 1368
QY 1988 ATGAATCCGTAATGATGTTGGGCTTACAGTGAAGATTAACAATATAGGAATCCACTA 2047
Db 1369 ATGAATCCGTAATGATGTTGGGCTTACAGTGAAGATTAACAATATAGGAATCCACTA 1428
QY 2048 TTGTGAGTTGTTTGTGAGAAATTCAAATGCTGATGACAAATGATATCTCCCTGACTAT 2107
Db 1429 TTGTGAGTTGTTTGTGAGAAATTCAAATGCTGATGACAAATGATATCTCCCTGACTAT 1488
QY 2108 GCAAAATTTGTGAAAACCTGTGTGCAAGGGTTGTCTTCTTACGTTTACAGACACCAAG 2167
Db 1489 GCAAAATTTGTGAAAACCTGTGTGCAAGGGTTGTCTTCTTACGTTTACAGACACCAAG 1548
QY 2168 ATAAAAATTTTAACTCGGGGACTACTATGATGATCTTATGTTTGAAGTTTACTTGAAA 2227
Db 1549 ATAAAAATTTTAACTCGGGGACTACTATGATGATCTTATGTTTGAAGTTTACTTGAAA 1608
QY 2228 GAGTGAAGTATGTTTCAAGGTTCTCTCTTATGCTGCTGCAACATATGAGCAAGATTGAG 2287
Db 1609 GAGTGAAGTATGTTTCAAGGTTCTCTCTTATGCTGCTGCAACATATGAGCAAGATTGAG 1668
QY 2288 CCGAGCATGTGAAAAGCTAGTGTCTATGACAGCACTGCAGAAAAGTTTTCCTTCCGCTATA 2347
Db 1669 CCGAGCATGTGAAAAGCTAGTGTCTATGACAGCACTGCAGAAAAGTTTTCCTTCCGCTATA 1728
QY 2348 CAGATTAAGAACTCGGCTGAAACCCAGAGATGTGCAAGAGACAGTGTTAAGTATGATCCTG 2407
Db 1729 CAGATTAAGAACTCGGCTGAAACCCAGAGATGTGCAAGAGAGTGTTAAGTATGATCCTG 1788
QY 2408 TTGGTAACTAATGTAGGCGGTGATGAGCTGGTGTCTTTATTTGCAAGACTGTAAAGAC 2467
Db 1789 TTGGTAACTAATGTAGGCGGTGATGAGCTGGTGTCTTTATTTGCAAGACTGTAAAGAC 1848
QY 2468 CTTCTGAAAACCTTTGAAACTATATGATTTATGCAATTTGAGCTGGGCTCTCAGAGATAGCG 2527
Db 1849 CTTCTGAAAACCTTTGAAACTATATGATTTATGCAATTTGAGCTGGGCTCTCAGAGATAGCG 1908
QY 2528 TTGATGAAGCTACTGTTGAAATGTCCGTTCTGATGTTTAAAGAGGCAAGTGTGAAGA 2587
Db 1909 TTGATGAAGCTACTGTTGAAATGTCCGTTCTGATGTTTAAAGAGGCAAGTGTGAAGA 1968
QY 2588 TCCTAGCTGCTGGGTGAGCAATTTGACTGATTTCACTGTTCAAGCAGAAAGTATTTTCTTA 2647
Db 1969 TCCTAGCTGCTGGGTGAGCAATTTGACTGATTTCACTGTTCAAGCAGAAAGTATTTTCTTA 2028
QY 2648 AAMGACGCTCATCTTTTCAAACGCAAGATATGTTTCTTCTATGAAATCTGATGTGCTA 2707
Db 2029 AAMGACGCTCATCTTTTCAAACGCAAGATATGTTTCTTCTATGAAATCTGATGTGCTA 2088
QY 2708 CCAATAGTATGATTAATGATGCAATTTTCAATATCTGCAATGCTCAAAAATATGCTTGT 2767
Db 2089 CCA----- 2091
QY 2768 TTTGTGAGTAAAGAACATATGTTCCACTATAATACATGTCCTCAAAAAGTTTATCAAGATTA 2827
Db 2092 ----- 2091
QY 2828 ACAAGTGTGAGTAAATTTCACTAATTAATGCTGCTGAAATTTTGTGATCAAACTGATAGA 2887
Db 2092 ----- 2091
QY 2888 CAGAAATGTAATTTTCACTCTCAACATTTTCTGTTTAAATTAACGTAGATTAAGATTTGC 2947

[illegible]

FEATURES	source
TITLE	Opereyong, R. M.
JOURNAL	Direct Submission
Submitted	(17-JAN-2003) Department of Plant Biology, Michigan State University, 166 Plant Biology Building, East Lansing, MI 48824, USA
location/Qualifiers	
1. .2438	/organism="Arabidopsis thaliana"
1. .2438	/mol_type="mRNA"
1. .2438	/db_xref="taxon:3702"
1. .2438	/ecotype="wasilewskija"
1. .2438	/gene="ARC6"
1. .2438	/note="synonym: At5g42480"
2. .2407	/gene="ARC6"
2. .2407	/codon_start=1
2. .2407	/product="divisin protein"
2. .2407	/protein_id="AA018646.1"
2. .2407	/db_xref="GI:33436354"
2. .2407	/translation="MEALSHVIGIGLSPOLCLRPATTKLRSHNSTTITCSAKWADRLSDENFTDSSSSSPATATTTATVLSPPSADREPHAPRIDEPQVLAGTFLPLDGRIRPEARVSRPQPGFSDDALISRRDLISRRREREDPVLASVEVVGSDLVTDPMDKVPKALCVLOEGEETVEIVRVEALKERLPSGPDVYVWALAFDLSRDMALEDPPDPTGYEVEERALKILDEEGASISLAPLRQIDTLEITETPRLYEILGLPGDDYAAKRLNGLSGVRITLINSVGGGASALVGLITREKPMNEAFILPTAAQVDLFPATNSIPAESFEVEVALVAQAETIGKPHLLODADQFQOLQDAKWAHEIPMLLYPTRNWMEIDFGLERGLCALIGKVDCEMMGLDSEDSQYRNPAIVEVVGSDDNNDLPGLCLLETWLAGVYFPRPRDKKPKLGDYVDDPMYLSYERVEVVGSDAAAATMARIGAEHVKASAMQALOKVPSRYTDRNSARPKVOETVSEVYGVNNGVDGEGCVPIAEKVRSENPETMDYAIRAGVSSSYDTETVENSVAEMLEKASVKILAAVVAIGILFSQKTFLEKSSSFQRKQWVSMESDVATIGSVRADSEALPRDARTKENTVVSUKQIKSLAFGPDHRIEMLPVLDRMLKITDRAETRAQGLVYDYLTLKLSVDSVTSASDGTALVEATLEESACSLDLVHPENNADVRYTTRVEVFWSGMKITGSGVLA"
misc_difference	474
misc_difference	/gene="ARC6"
misc_difference	/note="results in gly to glu substitution"
misc_difference	/replace="a"
misc_difference	1831
misc_difference	/gene="ARC6"
misc_difference	/note="results in glu to asp substitution"
misc_difference	/replace="c"
ORIGIN	
Query Match	47.8%; Score 1752.8; DB 15; Length 2438;
Best Local Similarity	79.3%; Pred. No. 0;
Matches 2436; Conservative	0; Mismatches 2; Indels 632; Gaps 5;
480	CATGAAGCTCTGAGTACGCTGGCATTTGCTCTCTCCCATTCCAATATATGCGGATTACC 539
1	CATGAAGCTCTGAGTACGCTGGCATTTGCTCTCTCCCATTCCAATATATGCGGATTACC 60
540	ACGGGCGACGACCAAAAGTCTCGAGGTAGCCCAACAACCTCTTACAACATATCTGCTCCGCGAG 599
61	ACGGGCGACGACCAAAAGTCTCGAGGTAGCCCAACAACCTCTTACAACATATCTGCTCCGCGAG 120
600	CAATGGGCGACGCTTCTCTCTCCGACTTCAATTTCACTCCGATTCCTCTCTCTCTC 659
121	CAATGGGCGACGCTTCTCTCTCCGACTTCAATTTCACTCCGATTCCTCTCTCTCTC 180
660	CTTGGCACCGGACCAACAACCGGCACTCTGCTCTCTGCGACCATCTATTGATCGTCC 719
181	CTTGGCACCGGACCAACAACCGGCACTCTGCTCTCTGCGACCATCTATTGATCGTCC 240
720	CGAAGCGCACGCTCCCATCTGATTTCTACACAGATATTAAGAGCTCAAAACATTT 779
241	CGAAGCGCACGCTCCCATCTGATTTCTACACAGATATTAAGAGCTCAAAACATTT 300
780	CTTAACCGAGGAATGAGAAGACATTCGAAGCTAGGATTTGAAACCGGCGGCAATTCGG 839
301	CTTAACCGAGGAATGAGAAGACATTCGAAGCTAGGATTTGAAACCGGCGGCAATTCGG 360
840	TTTCAGGACGAGCGCTTATTCAGCCGGAGACAGATTTCTCAAGCTGCTTGGCAAACTCT 899

|||||
Db 361 TTTGAGCAGACCGCTTAAATCAGCCGAGACAGATTCCTCAAGCTGCTCGAAACTCT 420
Oy 900 GTCTAACTCTGGGTAGAGAGATAGACAATGAAGTCTTCTGATATGAATGAAGAGCTAC 959
Db 421 GTCTAACTCTGGGTAGAGAGATAGACAATGAAGTCTTCTGATATGAATGAAGAGCTAC 480
Oy 960 AGTCATCATGATGTTCCCTGGGATGAAGTAATTTGATTTCCGAATATATAAGTTTCTT 1019
Db 481 AGTATCATCTATGTTCCCTGGGAT----- 505
Oy 1020 CGTTTAAATTCATGAATTTGGATTAAGGAAGAACTTTTATCTAGTGAAGTTCCTGGGG 1079
Db 506 -----AAGTTCTCTGGGG 518
Oy 1080 CTCTCTGATATTTGCAAGAGGTGGTGAAGTGAATGATGTTCTTGAGGTGGTGAAGCTC 1139
Db 519 CTCTCTGATATTTGCAAGAGGTGGTGAAGTGAATGATGTTCTTGAGGTGGTGAAGCTC 578
Oy 1140 TGCTTAAAGAGAGGTGGCTTAAGTCTGTTTAAAGCAAGATGTGGTTTAGTTATGGCGCTTG 1199
Db 579 TGCTTAAAGAGAGGTGGCTTAAGTCTGTTTAAAGCAAGATGTGGTTTAGTTATGGCGCTTG 638
Oy 1200 CGTTTCTGATATGTTCTGAGAGGATGCTATGAGCACTGATTTTATTAATGCTGTT 1259
Db 639 CGTTTCTGATATGTTCTGAGAGGATGCTATGAGCACTGATTTTATTAATGCTGTT 698
Oy 1260 ATGAGTTTGTGAGAGAGCTTTGAAGCTTTTACAGATGATTTGACTGCTTGGTAATTT 1319
Db 699 ATGAGTTTGTGAGAGAGCTTTGAAGCTTTTAC----- 731
Oy 1320 GACGAGCGTTGGCTTTATTAAGAACTTTCTTGATTTGATATCTTGTATTTGAGTCTTGTGT 1379
Db 732 ----- 731
Oy 1380 AGAGAGAAAGAGACAGTACGCTTGCACCGGATTTTCGTGCACAAATTTGATGAGACTTTGG 1439
Db 732 AGAGAGAAAGAGACAGTACGCTTGCACCGGATTTTCGTGCACAAATTTGATGAGACTTTGG 791
Oy 1440 AAGAGATCACTCCGGCTTATGCTTGAAGCTACCTTGAGCTTACCGCTTGAGTGAATGACG 1499
Db 792 AAGAGATCACTCCGGCTTATGCTTGAAGCTACCTTGAGCTTACCGCTTGAGTGAATGACG 851
Oy 1500 CTGCGAAAGAAGCTAAATGCTTTAAGCGGATGTGCGGAATATTTTGTGCTGTTGGAGAG 1559
Db 852 CTGCGAAAGAAGCTAAATGCTTTAAGCGGATGTGCGGAATATTTTGTGCTGTTGGAGAG 911
Oy 1560 GTGAGCAATCAGCTTGTGTGGGGGTTTGACCGGTGAAGATTATGATGAGCGCTTTT 1619
Db 912 GTGAGCAATCAGCTTGTGTGGGGGTTTGACCGGTGAAGATTATGATGAGCGCGTTTT 971
Oy 1620 TACGAATGACAGCTGCTGAGCAGTATACAGTTTGAATACCTTTTAAATTTCTTAGC 1679
Db 972 TACGAATGACAGCTGCTGAGC----- 992
Oy 1680 ATGATATTAACCTTAGGTTTCTCATTTTATATGATGTGTGTAGAGTTGATCTTTTGT 1739
Db 993 -----AGTTGATCTTTTGT 1008
Oy 1740 AGCTTACCCCAAGCAATATTTCCAGCAGATCATTTGAAGTTTACGAAGTTGCACTTGCTCT 1799
Db 1009 AGCTTACCCCAAGCAATATTTCCAGCAGATCATTTGAAGTTTACGAAGTTGCACTTGCTCT 1068
Oy 1800 TGTGCTCAAGCTTTTATTTGTTAAGAGCACAACCTTTTACAGATGCTGATTAAGCAATT 1859
Db 1069 TGTGCTCAAGCTTTTATTTGTTAAGAGCACAACCTTTTACAGATGCTGATTAAGCAATT 1128
Oy 1860 CCAGCACTTCAGCAGGCTAAGATATGCTATGAGAGATTCCTGCGATGTTGTATGATATAC 1919
Db 1129 CCAGCACTTCAGCAGGCTAAGATATGCTATGAGAGATTCCTGCGATGTTGTATGATATAC 1188
Oy 1920 ACGGAATATTTGGGATATGATCTTGGCTGTAGAAAGGGGACTCTGTGCACTGCTTATAGG 1979
|||||

Db 1189 ACGGAATATTTGGGATATGATCTTGGCTGTAGAAAGGGGACTGTGCACTGCTTATAGG 1248
Oy 1980 CAAAGTTGATGAATGCGGTATGTGTTGGCTTTAGACAGTGAAGGATTCACAAATATAGAA 2039
Db 1249 CAAAGTTGATGAATGCGGTATGTGTTGGCTTTAGACAGTGAAGGATTCACAAATATAGAA 1308
Oy 2040 TCCAGCTATTTGGAGTTTGTTTTGGAGAAATTCAAATCGTATGACAAATGATATCTCCC 2099
Db 1309 TCCAGCTATTTGGAGTTTGTTTTGGAGAAATTCAAATCGTATGACAAATGATATCTCCC 1368
Oy 2100 TGGACTATGCAAAATTTGTTGAAAACCTGTGGGAGGGGCTGTCTTCTAGTTCAAGAGA 2159
Db 1369 TGGACTATGCAAAATTTGTTGAAAACCTGTGGGAGGGGCTGTCTTCTAGTTCAAGAGA 1428
Oy 2160 CACCAAGATTAATAAATTTTAACTCGGGGACTACTATGATGATCTTATGTTTGAATTA 2219
Db 1429 CACCAAGATTAATAAATTTTAACTCGGGGACTACTATGATGATCTTATGTTTGAATTA 1488
Oy 2220 CTTGAAAAGATGAGGTAGTTCAAGGTTCTCTTTAGCTGTGCTGCAACTATGCAAG 2279
Db 1489 CTTGAAAAGATGAGGTAGTTCAAGGTTCTCTTTAGCTGTGCTGCAACTATGCAAG 1548
Oy 2280 GATTGGAGCCGAGATGTAAGAGCTAGTCTATGSCAGGCACTGCAAGATTTTCCCTC 2339
Db 1549 GATTGGAGCCGAGATGTAAGAGCTAGTCTATGSCAGGCACTGCAAGATTTTCCCTC 1608
Oy 2340 CCGCTATACAGATTAAGAACTCGGCTGAACCCAGAGATGTGCAGAGACAGTGTTTAGTGT 2399
Db 1609 CCGCTATACAGATTAAGAACTCGGCTGAACCCAGAGATGTGCAGAGACAGTGTTTAGTGT 1668
Oy 2400 AGATCTGTTGGTGAACAATGTAGGCGGTGATGTGAGCTGTGTCTTTATTTGCAAGC 2459
Db 1669 AGATCTGTTGGTGAACAATGTAGGCGGTGATGTGAGCCTGTGTCTTTATTTGCAAGC 1728
Oy 2460 TGTAAAGCCCTCGAAAACCTTTGAACCTAATGATTAATGCAATTTGACCTGGGCTGAGA 2519
Db 1729 TGTAAAGCCCTCGAAAACCTTTGAACCTAATGATTAATGCAATTTGACCTGGGCTGAGA 1788
Oy 2520 GAGTAGGCTGATGAACCTACTGTTGAATGTCCGTGCTGATATGTTAAAGAGGCAAG 2579
Db 1789 GAGTAGGCTGATGAACCTACTGTTGAATGTCCGTGCTGATATGTTAAAGAGGCAAG 1848
Oy 2580 TGTGAAGTCTTACCTGCTGTGTGTCGAATTTGACTATTTCACTGTTCAAGCCAGAACTA 2639
Db 1849 TGTGAAGTCTTACCTGCTGTGTGTCGAATTTGACTATTTCACTGTTCAAGCCAGAACTA 1908
Oy 2640 TTTTCTTAAAGAGCTCACTCTTTTCAACGAGAGATATGTTTCTCTATGAAATCTGA 2699
Db 1909 TTTTCTTAAAGAGCTCACTCTTTTCAACGAGAGATATGTTTCTCTATGAAATCTGA 1968
Oy 2700 TGTGCTAACCATAGATGATTAATATGATGCAATTTTCAATATCTGCAATTGCTCAAAAT 2759
Db 1969 TGTGCTAACCA----- 1979
Oy 2760 ATGCTGTTTTGTGAGCTTAAGAACATATGTTCCACTTAATATCATGTGCCAAAGTGTAC 2819
Db 1980 ----- 1979
Oy 2820 CAAGATTAACAAGTTGCTGAGTAAATTTCACTAATATATGCTTGAATTTTGTATCA 2879
Db 1980 ----- 1979
Oy 2880 ACTGTAGACAGAAATGTAAATTTCACTGCAACATTTCTGTTAAGAAATACATGATTA 2939
Db 1980 ----- 1979
Oy 2940 GAGATTCCTTATGTGTGACTTTGTCCAACTTTCTTCTTGATTTTCTTTCCATTT 2999
Db 1980 ----- 1979
Oy 3000 TAGGTCAGTCAGAGCTGACGATTCAGAAAGCACTTCCAGAAATGATGCTAAGAGCTGAG 3059
Db 1980 TAGGTCAGTCAGAGCTGACGATTCAGAAAGCACTTCCAGAAATGATGCTAAGAGCTGAG 2039
|||||

Db	121	AAATGGCGCCACCGCTTCTCTCCGACTTCAATTTACCTTCGATTTCTCTCTCTCTCC	180
Oy	661	TTGCGCACCGCACCCACCGCACCTCTCGTCTCTGCGCACCATCTAATGATGATGCC	720
Db	181	TTGCGCACCGCACCCACCGCACCTCTCGTCTCTGCGCACCATCTAATGATGATGCC	240
Oy	721	GAAGCGCACGTCGCCATTCGCCATGATTTCTTACCAAGGATTTAGAGCTCAAAACATTTG	780
Db	241	GAAGCGCACGTCGCCATTCGCCATGATTTCTTACCAAGGATTTAGAGCTCAAAACATTTG	300
Oy	781	TTAACCGATGGAATCAGAAAGACATTCGAGCTGAGGGTTGAAACCGCGCAATTCGGT	840
Db	301	TTAACCGATGGAATCAGAAAGACATTCGAGCTGAGGGTTGAAACCGCGCAATTCGGT	360
Oy	841	TTACGCGCAGACGCTTTAATCAGCGCGAGACAGATTTCTTCAAGCTGCTTGCGAAACTCTG	900
Db	361	TTACGCGCAGACGCTTTAATCAGCGCGAGACAGATTTCTTCAAGCTGCTTGCGAAACTCTG	420
Oy	901	TCATAATCTCGGCTTGAAGAAGAGTACATAGAGTCTTCTTGATGATGAAGAGCTACA	960
Db	421	TCATAATCTCGGCTTGAAGAAGAGTACATAGAGTCTTCTTGATGATGAAGAGCTACA	480
Oy	961	GTCATCATGATGATGCTTGGGATTAAGGTAATTCGATTTGCGAATAATAAGTTTCTTC	1020
Db	481	GTCATCATGATGATGCTTGGGATTAAGGTAATTCGATTTGCGAATAATAAGTTTCTTC	504
Oy	1021	GTTTAAATTTTATGAAATTTGATTAAGAAAGAACTTTATCTAGTGAAGGTTCTTGAGGGC	1080
Db	505	-----AAGTCTCTGGATGC	518
Oy	1081	TCTCTGTATATTCAGAAAGGTGCTGAGACTGAGATAGTCTTTCGGGTGCTGAGGCTCT	1140
Db	519	TCTCTGTATATTCAGAAAGGTGCTGAGACTGAGATAGTCTTTCGGGTGCTGAGGCTCT	578
Oy	1141	GCTTAAGGAGAGGTTCCTAAGCTGTTTAACCAAGATGCTGTTTATGATATGAGGCTGCG	1200
Db	579	GCTTAAGGAGAGGTTCCTAAGCTGTTTAAACCAAGATGCTGTTTATGATATGAGGCTGCG	638
Oy	1201	GTTTCTCGATGCTTCGAGAGGATGCTATGAGATGAGATCACTGATTTTATTACTGCTTA	1260
Db	639	GTTTCTCGATGCTTCGAGAGGATGCTATGAGATGAGATCACTGATTTTATTACTGCTTA	698
Oy	1261	TGAGTTGTGAGAAAGCTTTGAAGCTTTTACAGGTAGTGTGAACTTCCTTGTAATTTGG	1320
Db	699	TGAGTTGTGAGAAAGCTTTGAAGCTTTTAC	730
Oy	1321	ACGAGCGTGGCTTTATAAGAACTTCTTGATTTGATTAATCTTGATTTGACTTGTGTGA	1380
Db	731	-----A	731
Oy	1381	GGAGGAAGAGCAAGTAGACTCTTGACCGGATTTACGTGACCAATATGATGAGACTTTGGA	1440
Db	732	GGAGGAAGAGCAAGTAGACTCTTGACCGGATTTACGTGACCAATATGATGAGACTTTGGA	791
Oy	1441	AGAATCATCTCCGCGTTATGCTCTTGAGACTTTCGCTTACCGCTTGATGATTAATTCGC	1500
Db	792	AGAATCATCTCCGCGTTATGCTCTTGAGACTTTCGCTTACCGCTTGATGATTAATTCGC	851
Oy	1501	TGCGAAAAGACTAAATGCTTTAAGCGGTGTCGGGAATATTTTGTGTGCTGTGAGAGAGG	1560
Db	852	TGCGAAAAGACTAAATGCTTTAAGCGGTGTCGGGAATATTTTGTGTGCTGTGAGAGAGG	911
Oy	1561	TGAGACATCAGCTCTGTTGGGGGTTTGACCGGTGAAGAGTTTATGATGAGAGGCGTTT	1620
Db	912	TGAGACATCAGCTCTGTTGGGGGTTTGACCGGTGAAGAGTTTATGATGAGAGGCGTTT	971
Oy	1621	ACGATATGACAGCTGCTGAGCAGGTATACAGTTTATGATTAATTTTAAATTTCTTTAGCA	1680
Db	972	ACGATATGACAGCTGCTGAGC-----	991
Oy	1681	TGATATTAACCTTAAAGTTCCTCATTTTAAATGATGTTGTGTGATGAGTTGATCTTTTGTGA	1740

Db	992	-----AGGTGATCTTTTGTGA	1008
QY	1741	GCTACCCCAAGCAATATTC	1800
Db	1009	GCTACCCCAAGCAATATTC	1068
QY	1801	GTGGCTCAAGCTTTTAT	1860
Db	1069	GTGGCTCAAGCTTTTAT	1128
QY	1861	CAGCAACTTCAGCAGGCT	1920
Db	1129	CAGCAACTTCAGCAGGCT	1188
QY	1921	CGGAATTAATGGAGAT	1980
Db	1189	CGGAATTAATGGAGAT	1248
QY	1981	AAAGTTGATGAATGCC	2040
Db	1249	AAAGTTGATGAATGCC	1308
QY	2041	CCAGCTATTGCGAGTT	2100
Db	1309	CCAGCTATTGCGAGTT	1368
QY	2101	GGACTATGCAAAAT	2160
Db	1369	GGACTATGCAAAAT	1428
QY	2161	ACCAAAAGTAAAAA	2220
Db	1429	ACCAAAAGTAAAAA	1488
QY	2221	TTGGAAGAAGTGAAG	2280
Db	1489	TTGGAAGAAGTGAAG	1548
QY	2281	ATTGAGCCGAGCAT	2340
Db	1549	ATTGAGCCGAGCAT	1608
QY	2341	CGCTATACAGATGA	2400
Db	1609	CGCTATACAGATGA	1668
QY	2401	GATCCGTGTGGTA	2460
Db	1669	GATCCGTGTGGTA	1728
QY	2461	GTAAAGACCTCTGA	2520
Db	1729	GTAAAGACCTCTGA	1788
QY	2521	AGTAGCGTTGATGA	2580
Db	1789	AGTAGCGTTGATGA	1848
QY	2581	GTGAAGATCCATG	2640
Db	1849	GTGAAGATCCATG	1908
QY	2641	TTTCTTAAGAAGC	2700
Db	1909	TTTCTTAAGAAGC	1968
QY	2701	GTGCGTACATAG	2760
Db	1969	GTGCGTACATAG	1978
QY	2761	TGCTTGTTTTGT	2820
Db	1979	-----	1978


```
Db 133167 TCCTGCATATGAAAGCTGCCGAGACAGTGCTCATAGAAATATATGCAATGTTTTTAT 133226
Qy 1671 TCTTAGCATGATATAACTTAGGTTCTCATTTTAATGAT----- 1713
Db 133227 TTAATTTTCTCCGCCCCCTTGATTTTGTATGAAAGTTTACATTTATCTTTCA 133286
Qy 1714 -----GTTGTGTAGTGTGATCTTTTGTAGTACCCCAAG 1751
Db 133287 TTAATCTTTTATCTCCCAATCTTTTTTTTAAAGTCGAACCTTTTGTAGCAACCTAG 133346
Qy 1752 CAATATTCAGCAGACAGTCAATTGAAAGTTTACGAAGTTGACCTTGTCTTGTGGCTCAAGC 1811
Db 133347 CAATATTCAGCAGTGAAGTTTGTAGGCCCTATGCGCTTGCACCTTGATGGTTGCAAGC 133406
Qy 1812 TTTTATGTAGTAAGCCACACCTTTTACAGAGTGTGATTAACAATTCCAGCACTTCA 1871
Db 133407 CTTTGTAGTTAAAGCCCACTTTATCCAGATGTGTATTAATTTATCCATCAACTTCA 133466
Qy 1872 GCAGGCTAAGTAATGCGCTATGAGAGATTCCTGCGATGTTGTATGATAC-----ACGGA 1925
Db 133467 ACAAACTAAGTAACAATAATAGAGATGCTCCTCTGTTTATCTCCCATGAGATGGA 133526
Qy 1926 TAAATGGAGATAGACTTGGCTGTAGAAAGGGACCTGTGTGACCTGCTTATAGCCAAAT 1985
Db 133527 GAAAGAGAGAGTGAATTTGACATTAGAAAGGGGCTGTGTGACCTGCTTGTGGGAGCT 133586
Qy 1986 TGATGAATGCCGTATGTGGTGGCTTAGACAGTAGAGATTCACAAATATAGGATCCAGC 2045
Db 133587 TGATCAATGTCAATCATGTGTGGTCTCGATAGTGAAGCAGTCACTTATCGAAACCTATC 133646
Qy 2046 TATTTGTGAGTTTGTTTGTGAGAAATTCAAATCGTATGACAAATATGATCTCTGAGCT 2105
Db 133647 TATTAATGACTTTAATATGAAAACGAAAAGGTATGAAAGCAGTGAATCTTCTGAGCT 133706
Qy 2106 ATGCAATTTGTGGAACCTGTTGGCAGGGTGTCTTCTTCTAGATTCAAGACACCA 2165
Db 133707 CTGTAATTTGTGAGACATGGTGTATGAGAGGTGTTTCCCTAGTTTGAATATCTTA 133766
Qy 2166 AGATAAAAAATTT 2178
Db 133767 AGAGCAAACTTT 133779
```

RESULT 8
AP008208_011/c

WPCOMMENT

Sequence split into 360 fragments. LOCUS AP008208 Accession AP008208

Fragment Name	Begin	End
AP008208_000	1	110000
AP008208_001	100001	210000
AP008208_002	200001	310000
AP008208_003	300001	410000
AP008208_004	400001	510000
AP008208_005	500001	610000
AP008208_006	600001	710000
AP008208_007	700001	810000
AP008208_008	800001	910000
AP008208_009	900001	1010000
AP008208_010	1000001	1110000
AP008208_011	1100001	1210000
AP008208_012	1200001	1310000
AP008208_013	1300001	1410000
AP008208_014	1400001	1510000
AP008208_015	1500001	1610000
AP008208_016	1600001	1710000
AP008208_017	1700001	1810000
AP008208_018	1800001	1910000
AP008208_019	1900001	2010000
AP008208_020	2000001	2110000
AP008208_021	2100001	2210000
AP008208_022	2200001	2310000
AP008208_023	2300001	2410000
AP008208_024	2400001	2510000

AP008208_025	2500001	2610000
AP008208_026	2600001	2710000
AP008208_027	2700001	2810000
AP008208_028	2800001	2910000
AP008208_029	2900001	3010000
AP008208_030	3000001	3110000
AP008208_031	3100001	3210000
AP008208_032	3200001	3310000
AP008208_033	3300001	3410000
AP008208_034	3400001	3510000
AP008208_035	3500001	3610000
AP008208_036	3600001	3710000
AP008208_037	3700001	3810000
AP008208_038	3800001	3910000
AP008208_039	3900001	4010000
AP008208_040	4000001	4110000
AP008208_041	4100001	4210000
AP008208_042	4200001	4310000
AP008208_043	4300001	4410000
AP008208_044	4400001	4510000
AP008208_045	4500001	4610000
AP008208_046	4600001	4710000
AP008208_047	4700001	4810000
AP008208_048	4800001	4910000
AP008208_049	4900001	5010000
AP008208_050	5000001	5110000
AP008208_051	5100001	5210000
AP008208_052	5200001	5310000
AP008208_053	5300001	5410000
AP008208_054	5400001	5510000
AP008208_055	5500001	5610000
AP008208_056	5600001	5710000
AP008208_057	5700001	5810000
AP008208_058	5800001	5910000
AP008208_059	5900001	6010000
AP008208_060	6000001	6110000
AP008208_061	6100001	6210000
AP008208_062	6200001	6310000
AP008208_063	6300001	6410000
AP008208_064	6400001	6510000
AP008208_065	6500001	6610000
AP008208_066	6600001	6710000
AP008208_067	6700001	6810000
AP008208_068	6800001	6910000
AP008208_069	6900001	7010000
AP008208_070	7000001	7110000
AP008208_071	7100001	7210000
AP008208_072	7200001	7310000
AP008208_073	7300001	7410000
AP008208_074	7400001	7510000
AP008208_075	7500001	7610000
AP008208_076	7600001	7710000
AP008208_077	7700001	7810000
AP008208_078	7800001	7910000
AP008208_079	7900001	8010000
AP008208_080	8000001	8110000
AP008208_081	8100001	8210000
AP008208_082	8200001	8310000
AP008208_083	8300001	8410000
AP008208_084	8400001	8510000
AP008208_085	8500001	8610000
AP008208_086	8600001	8710000
AP008208_087	8700001	8810000
AP008208_088	8800001	8910000
AP008208_089	8900001	9010000
AP008208_090	9000001	9110000
AP008208_091	9100001	9210000
AP008208_092	9200001	9310000
AP008208_093	9300001	9410000
AP008208_094	9400001	9510000
AP008208_095	9500001	9610000
AP008208_096	9600001	9710000
AP008208_097	9700001	9810000

AP008208_098	9800001	9910000
AP008208_099	9900001	
AP008208_100	1000001	10010000
AP008208_101	1010001	10210000
AP008208_102	1020001	10310000
AP008208_103	1030001	10410000
AP008208_104	1040001	10510000
AP008208_105	1050001	10610000
AP008208_106	1060001	10710000
AP008208_107	1070001	10810000
AP008208_108	1080001	10910000
AP008208_109	1090001	11010000
AP008208_110	1100001	11110000
AP008208_111	1110001	11210000
AP008208_112	1120001	11310000
AP008208_113	1130001	11410000
AP008208_114	1140001	11510000
AP008208_115	1150001	11610000
AP008208_116	1160001	11710000
AP008208_117	1170001	11810000
AP008208_118	1180001	11910000
AP008208_119	1190001	12010000
AP008208_120	1200001	12110000
AP008208_121	1210001	12210000
AP008208_122	1220001	12310000
AP008208_123	1230001	12410000
AP008208_124	1240001	12510000
AP008208_125	1250001	12610000
AP008208_126	1260001	12710000
AP008208_127	1270001	12810000
AP008208_128	1280001	12910000
AP008208_129	1290001	13010000
AP008208_130	1300001	13110000
AP008208_131	1310001	13210000
AP008208_132	1320001	13310000
AP008208_133	1330001	13410000
AP008208_134	1340001	13510000
AP008208_135	1350001	13610000
AP008208_136	1360001	13710000
AP008208_137	1370001	13810000
AP008208_138	1380001	13910000
AP008208_139	1390001	14010000
AP008208_140	1400001	14110000
AP008208_141	1410001	14210000
AP008208_142	1420001	14310000
AP008208_143	1430001	14410000
AP008208_144	1440001	14510000
AP008208_145	1450001	14610000
AP008208_146	1460001	14710000
AP008208_147	1470001	14810000
AP008208_148	1480001	14910000
AP008208_149	1490001	15010000
AP008208_150	1500001	15110000
AP008208_151	1510001	15210000
AP008208_152	1520001	15310000
AP008208_153	1530001	15410000
AP008208_154	1540001	15510000
AP008208_155	1550001	15610000
AP008208_156	1560001	15710000
AP008208_157	1570001	15810000
AP008208_158	1580001	15910000
AP008208_159	1590001	16010000
AP008208_160	1600001	16110000
AP008208_161	1610001	16210000
AP008208_162	1620001	16310000
AP008208_163	1630001	16410000
AP008208_164	1640001	16510000
AP008208_165	1650001	16610000
AP008208_166	1660001	16710000
AP008208_167	1670001	16810000
AP008208_168	1680001	16910000
AP008208_169	1690001	17010000
AP008208_170	1700001	17110000

	Query Match	Best Local Similarity	8.7%;	Score 317.6;	DB 15;	Length 110000;
	Matches 1459;	Conservative	0;	Mismatches 1259;	Indels 202;	Gaps 16;
QY	744	TCGATTTCTACAGGATATTAGAGCTCAACACATTTCTTAACCGATGAAATCAGAGAGC	803			
DB	81995	TGGCTTTGCTCAGGTTCTACAGGGGCGAGCCACATTTCTTGCGATGACATCAGAGAGGC	81933			
QY	804	ATTGCAAGCTAAGGGTTTCGAAACCGCGCAATTGCTTACGAGAGAGAGCGTTTAATCAG	863			
DB	81935	GTTTCGAGGCAAGATVAGCCAGGCCACCGCATGTGCTACAGCACGAGATCTCTTGTGG	8187			
QY	864	CCGGAGCAGATTTCTTCAGCTGCTTTCGAACTGTCTTAATCTTCGCTAGAAAGA	923			
DB	81875	TCGTGCACAATGCTGAGATTGCCATGACATCTCATGAAACAGAACTCCGCACTCA	8181			
QY	924	GTAACAATGAAGTCTTCTTGATGATGAAGAAGCTACGTCATCACTGATGTTCCCTGGGA	983			
DB	81815	GTAATGATCGCGCTTCTGAGAACCGTGAAGAAGCTCTCACATGATATTGCTTGGGA	8175			
QY	984	TAAAGTAATTCGATTTCCGAAATATTAAGTTTCT-----TCGT	1022			
DB	81755	CAAGGTGAGGCTTTAAGGGGCGACCAAGGGGTGTGTGGATCTTGTCAAATAGCACTAAT	8169			
QY	1023	TTTAATTTTCAATGAATTGATTAAGAAAGAACTTTATCTAATGAAGGTTCTGGGGCTC	1082			
DB	81695	GTAATTAATCTCTCTCTGCTTAATTAATTAATTTGGGTCCTTCTTGAAAGCAGGTTCCAGGTGT	8163			
QY	1083	TCTGTGTAATTCGAAGAAGTGTGTGACGTGAGTAGTCTTTCGAGTGTGTGAGGCTTCGC	1142			
DB	81635	TGTGTGGCTTTCAGAGAGCTGGGGAAGCACTTGTGTGCTTTGAATCTGGAACAGATTGC	8157			
QY	1143	TTAAGGAGAGGTTGCTTAAGTCTTTAAGCAAGATGTGTTTAACTTAATGCGCTTGCGT	1202			
DB	81575	TTCTGATTCGCGCACCCMACCGCTTCAAGAGAGCGTGGCTAGAGCATGCTCTGCTT	8151			
QY	1203	TTCTCATGTCTCGAGGAGATGCTATGAGCATTTGATCCACTGATTTTAACTAGGTTATG	1262			
DB	81515	ATGTGATCTATCAAGAGGATGCTATGAGCAAGCCCTCCAGATGTAATTGGCTGCTGCG	8145			
QY	1263	AGTTTGTGAAGAGCTTTGAAGCTTTTACAG-----TAG	1298			
DB	81455	AGGTGCTCGAAGAGGCTCTCAAGCTCTTGAGAGTGAATGTTGCTTTTAACTCCCTAC	8139			
QY	1299	TTTGACTGTGTTTGTGAATTTGACGAGCGTGGCTTTATPAAGACTTCTTAATTTG---	1355			

Db 81395 TTTTTCATTGGGCTTATATTTGAATAGATAAATTAATTAACCTGTGAAGATTGCTTT 81336
Qy 1356 -----ATACTTGTATTTAGTCTTG----- 1376
Db 81335 GGGACACATGGGCTATATTTTATTAATAAGATGTGATTTCTTGAAACATATTAATA 81276
Qy 1377 -----TTAGAGAGAGAGAGACATAGTACCTTGACCGGATTTAGGTGACAAA 1424
Db 81275 GGAACCTTACATCTAGAGAGAGATGAGACAACTGACCTGATCTGCTTTTCACAGA 81216
Qy 1425 TTGATAGAGCTTTGGAGAGAGATCACCCGCTTATGTCTGGAGCTACCTGGCTTACCGC 1484
Db 81215 TTGATGAACCTTCAGAGAGATTAACCTCGCTGTGTATTTGAGGCTTCTCCCTTCTTA 81156
Qy 1485 TTGTGATGATTAACCTGCGAAGAACTAATATGTTTAAACCGTGTGCGAATATTTGT 1544
Db 81155 TTGACACAGAGCATATAGAAGCGCCAAAGAGGCTTCAAGTTCGAGAAACATTTTGT 81096
Qy 1545 GGTCTGTTGAGAGAGGTGAGCATCACTCTTGTGGGGGTTGACCCGTGAAAGTTTA 1604
Db 81095 GGAGCGTTGGCAGAGAGGATTTGCTACCGTTGAGAGAGATTTCTGCTGAAGCCCTTCA 81036
Qy 1605 TGAATGAGCGCTTTTACGAATGACAGCTGCTGACAGGT-----ATACAGTTTNG 1655
Db 81035 TGAACGAGGCTTTTGAAGATGACATCAATTTGAACAGGCTTCGATCGATGTTTGTAT 80976
Qy 1656 ATACCTTTTAAATTTCTTTAGCATGATATTAATCTTAAAGTTCATTTTAAATGATGT 1715
Db 80975 GAAAGAGCTGCTATTTATTTAGCAGAGACATCTTTGAAACATGTTACTCCCTTTC 80916
Qy 1716 TGT-----GTGTAGGTGATCTTTTGTAGCTACCCACAGCAATATTCAGCAGATCA 1770
Db 80915 TGTTCACCTCCCAATGATGATTTCTTTCAAAAACCGAATAGACTTCCTCTGAAAG 80856
Qy 1771 TTTGAAGTTTACGAAGTTGCACTGTCTGTGAGCTTAAAGCTTTATTTGTAAGAGCA 1830
Db 80855 TTTGAAATTTACATGTAGACATTTGCAATGTGCTGACAGCATTTATTAATTAAGCAAG 80796
Qy 1831 CACCTTTTACAGAGTCTGATTAAGCAATTCAGCACTTCAGCAGGCTAAGGTAAATGCT 1890
Db 80795 CAATTCATCATGATGCGGATGATCTTTTGAACAACCTCAGAAATTCACATAGGTTCT 80736
Qy 1891 ATGAGATTTCTGCGATGTTGTATGATACACGAAATTAATGGGAATGACTTGGCTCTA 1950
Db 80735 CATTTATGCTTATG-----ATATGAGATGACCTTGCATTTG 80700
Qy 1951 GAAAGGGGACCTGTGCACTGCTTAPAGGCANAAGTTGATGATGCGTATGCTTGGGC 2010
Db 80699 GAAAGGGCATTTCTGCTCATTTGCTAGTCGAGATGTTAGCAAGTGAATGTGGCTTGA 80640
Qy 2011 TTAGAAGTGAAGATTCACAAATATAGAAATCCAGCTATTTGTGAGTGTGTTGAGAT 2070
Db 80639 ATGTAATATGAGCTTCCACATACAGAGACCCCAAAATTTCTAGATTTATTTGTGACAA 80580
Qy 2071 TC---AAATCGTATGACATGATGATCTCCCTGACATATGCAAAATTTGTTGAAACCTGG 2127
Db 80579 TCTAGCATCAAGTGAAGAGATGATCTTCTCCAGGCGTGTGCAACCTTTTGGAGACTTGG 80520
Qy 2128 TTGGAGAGGGGTGCTTTCCTAGTTCAGAGACACAAAGATAAAATTTAAACTCGGG 2187
Db 80519 CTTATCTTTAGAGTTCCTTCTAGAGACAGAGATCTCGGGGCTACAGTTCAGACTTGA 80460
Qy 2188 GACTACTATGATGATCTTATGTTTGAAGTACTTGGAAAGAGTGAAGTATTCAGAGT 2247
Db 80459 GATTACTAGATGATTCGAAATTTTAACTTACCTTAGAAGATGAGAGGTGTGTGCT 80400
Qy 2248 TCTCTTTAGCTGCTGCTGCAACTATGSCAAGATTTGAGC-----C 2289
Db 80399 TCTCATTTGGCTGCTGCTGCTCTATTTGCAAACTTGGGTCTCAAGCTACAGCTGACACT 80340
Qy 2290 GAGCAGATGGAAGCTAGTGTATGACAGGCACTGCAAGAAAGTTTTCCTCCGCTATATCA 2349
Db 80339 GGTACTGTGAATTCAAATGCTATTCAGAGGCTTCCACAGGTTTTTTCATTGATAGACAG 80280

Qy 2350 GATGAAACTGCGCTGAACCCAGAGATGTCAGAGACAGTGTTTAGTATGATCTGTT 2409
Db 80279 TTAGACAGTGCAGCCATGAAAAATCTAAAGATGCGCTGGGGATATCTTGAATAATTTT 80220
Qy 2410 GGTAAATATGAGGCCGCTGATGTGTGAGCTGG-----TGCTTTATTT 2451
Db 80219 GACCAGAAAAATGACCTGCTCATGATTCAGAAATGCCCTTTGAAATATATCTGTGCT 80160
Qy 2452 GCAGAACCTGTAAGACCTCTGAAAACCTTGAACCTATGATTAATGCAATTCAGCTGGG 2511
Db 80159 GGCGCATCTGTTGACCTGTTGGCAGTAAATTTGGGGCCAAATATTTGCCCTCGTAAGAGCCC 80100
Qy 2512 GTCTCAGAGATGACGTTGATGAAACTACTGTTGAAATGTCCGTTGCTGATATGTTAAAG 2571
Db 80099 CTTTCTGCTATTTAGAGATGAGCATGATCTGTGGCAGTGTCTAATATGTCGACTACT 80040
Qy 2572 GAGCAGATGTAAGATTCCTAGCTGCTGTGTGTGCAATGGAATGATTTCACTGTTACG 2631
Db 80039 GATGTGAGTAAATGCTTTTCATT--TGCTAAAGTACTCAACACACTGAAATTTACTG 79982
Qy 2632 CAGAAATATTTTCTTAAAGAGCTCATCTTTTCAACGCAAGATATGTTTCTTATG 2691
Db 79981 CTCAGCTGTTTACTGGGATTCCTTGTGGCTGATTAAGATATGTTCTTGATCTTATAGCT 79922
Qy 2692 GAATCTGATGCTGCTACCATAGGATGATTAATATGATGCAATTTTCATATATCTGACTG 2751
Db 79921 TTGCATTAATGTTTAAACAAAGCTATACATCAAAATATTTTGTCTCATACATACACATG 79862
Qy 2752 CTCAAATATGCTGTTTGTGTGAGCTAAGAACATATGTTCCACTTAATATACATGCCAA 2811
Db 79861 CTACTAATATTCGTAAATGATATATCAATAGAGATTAATTTTGATGATGATATGTTAAG 79802
Qy 2812 AGTTGTACCAAGATTAACAAGTTGCTGAGTAAATTTCACTAATATGCTGCTTG----A 2866
Db 79801 ATCTGCATAGATGAAAGATTCAGTACATATTCCTAATATGATGCAAGCTGGCAGAAAG 79742
Qy 2867 ATTTTGTATCAACCTGTAGACAGAAATGTAATTTTCACTCTCAACTTTCTGTTTGA 2926
Db 79741 ATATTTGTCGCAAGTGCAGAGTATCAAAATCTAAGGCTTTGGGACCAACATTCGGTTG 79682
Qy 2927 TAACTGAGATTTAGAAATTCGCTTATGCTGCTTGTGCCAATTTTCTTTC--CTGTATTT 2985
Db 79681 CATCATTTGCAAGAGTATGTTTAAACTTTGACTTTTCAAGCTTCTTACCTGGCCACTG 79622
Qy 2986 TTTTCTTTTGCATTTAAGGTCAGTCAAGCTGACGATTCAGAAACACTTCCAGAAATGA 3045
Db 79621 CTCTTACCTATACCTTACGCTTCTTATAGAGATTAATTCAGATTAATCTCTCTATAGGA 79562
Qy 3046 TGTAGAGCTGCAAGAAATATATATTCACAAATGSCAGAAATTAATGTCTGCTTTTGG 3105
Db 79561 TATATGTCACGCAATGTACTTTTACAAATGTTGTTTACAAATATCTTTCT----- 79511
Qy 3106 GCTGTATCCGCGCATAGAAATGTTTACAGAGCTGAGGGAATTAATCTACAAATTCATCA 3165
Db 79510 ---TGTTTGAATCAACACATGCTTCAATGCACTGAACAAATTTTCAAGAACTATATGC 79454
Qy 3166 TTGTGTAAAACTGTTGACATGATTAATGATGCTGCTGCTTGTGATTT---CTGTATTT 3222
Db 79453 CACTGCAATGATATCCCGTATGATGAGCGGTCAAAATCTGTATGATTTTCCATGTTCT 79394
Qy 3223 TATAGTTTTGATGAGCGAATGCTGAAGATTTGGAATGACAGACGCTGAAACTGCGC 3282
Db 79393 TATAGTTTCTTGAATGCAACATGTCTAAAGGTGTGATGACCGGACGCGAGATTTGAGC 79334
Qy 3283 AGCTTGGGTTGTTATGATTTATCACTGTTGAAACATATCTGTTGACAGTGTGCACTCT 3342
Db 79333 GTCAATGGGTGTTTGTGGAATATCACTATTCGATGGAAGATTAAGATCACTATCT 79274
Qy 3343 CAGCAGATGGAACCCGATCTGTGTGGAAGCACTCTGAGAGATGCTGTGTCTATCTG 3402
Db 79273 CCTATGATGTGACGAGCCACTGTGAGAGGCTACGATTTGATGAGGCGCAACTTACTG 79214

QY 3403 ATTGGTTCATCCAGAAAAGTCTAGTGTGAAAGCTACACAGAGATGAG 3462
 DB 79213 ATGTTACTAGCCCAAGAAATGATTCATGACCAAAATATACCTACCGGTATGAGA 79154
 QY 3463 TTTTCTGCTCCAG--TAAAGGTGAAATATACAGAGGCTCTGTTCTTGATCATAT 3519
 DB 79153 TGGCTCTTCCAGAGCTAGAGGGGTGAGATTAACGAAAGACAGATCTCAAGTCGTAGA 79094
 QY 3520 ATACTCATATGTAGCATGTC 3539
 DB 79093 AAGAGTGTGACATGACTTGTTC 79074
 RESULT 9
 LOCUS AP004885/c 150462 bp DNA linear PLN 15-SEP-2004
 DEFINITION Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 2,
 PAC clone: P0575F10.
 ACCESSION AP004885
 VERSION AP004885.3 GI:41053009
 SOURCE
 ORGANISM
 Oryza sativa (japonica cultivar-group)
 Oryza sativa (japonica cultivar-group)
 Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzeae; Oryza.
 REFERENCE
 AUTHORS 1 Sasaki, T., Matsumoto, T. and Yamamoto, K.
 TITLE Oryza sativa nipponbare (GA3) genomic DNA, chromosome 2, PAC
 clone: P0575F10
 JOURNAL Published Only in Database (2002)
 REFERENCE 2 (bases 1 to 150462)
 AUTHORS Sasaki, T., Matsumoto, T. and Yamamoto, K.
 TITLE Direct Submission
 JOURNAL Submitted (20-MAR-2002) Takuji Sasaki, National Institute of
 Agrobiological Sciences, Rice Genome Research Program, Kannondai
 2-1-2, Tsukuba, Ibaraki 305-8602, Japan
 (E-mail: tsasaki@nias.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/
 Tel:81-298-38-7441, Fax:81-298-38-7468)
 On Jan 21, 2004 this sequence version replaced gi:18142429.
 Genes were predicted from the integrated results of the following:
 GENSCAN (http://CCR-081.mit.edu/GENSCAN.html), GENESH
 (http://www.softberry.com/), GeneMark.hmm
 (http://opal.biology.gatech.edu/GeneMark/), GlimmerM
 (http://www.tigr.org/db/glimmer/glmr_form.html), RiceHMM
 (http://rgp.dna.affrc.go.jp/RiceHMM/), SplICEPredictor
 (http://bioinformatics.lasstate.edu/cgi-bin/sp.cgi), sim4
 (http://www.tigr.org/software/glimmer/), BLASTN and BLASTX. The
 genomic sequence was searched against NCBI Nonredundant Protein
 database, nr (ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA
 sequence database at RGP or DBJ. Protein homologues of the coding
 regions were searched against NCBI Nonredundant Protein database
 with BLASTP. ESTs represent the identified cDNA sequences using
 BLASTN with the corresponding DBJ accession no. and RGP clone ID.
 Full-length cDNAs represent the identified cDNA sequences using
 BLASTN with the corresponding DBJ accession no.
 A gene with identity or significant homology to a protein is
 classified based on the protein name to indicate the homology level
 such as same name, 'putative-' and '-like protein'. A gene without
 significant homology to any protein but with full-length cDNA or
 EST homology (covering almost the entire length of partial
 sequence) is classified as an 'unknown' protein. A gene predicted
 by two or more gene prediction programs is classified as a
 'hypothetical' protein according to IRGSP standard. A gene
 predicted by a single gene prediction program is also classified as a
 'hypothetical' protein and is included as a
 miscellaneous feature of the sequence.
 The orientation of the sequence is from T7 to SP6 of the PAC clone.
 This sequence of P0575F10 clone has an overlap with G11020 C02
 (DBJ: AP004078) clone at 5' end and with P0482P12 (DBJ: AP005111)
 clone at 3' end. Detailed information on overlap and assembly
 quality together with annotation of this entry is available at

FEATURES
 source
 http://rgp.dna.affrc.go.jp/GenomesSeq.html.
 Location/Qualifiers
 1..150462
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="genomic DNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /chromosome="2"
 /clone="P0575F10"
 complement(4987..8112)
 /gene="P0575F10.1"
 complement(join(4987..5412,5520..5679,5764..5914,
 6096..6247,6738..6859,7864..8112))
 /gene="P0575F10.1"
 /note="supported by full-length cDNA(s): AK106553"
 complement(join(5003..5412,5520..5679,5764..5914,
 6096..6247,6738..6859,7864..8084))
 /gene="P0575F10.1"
 /note="supported by full-length cDNA(s): AK073059"
 complement(join(5315..5412,5520..5679,5764..5914,
 6096..6247,6738..6859,7864..7876))
 /gene="P0575F10.1"
 /note="contains EST(s): AU070160(BE1534)
 contains full-length cDNA(s): AK106553,AK073059"
 /codon_start=1
 /product="putative eukaryotic translation initiation
 factor 3, subunit 1 alpha, 35kDa"
 /protein_id="BAD07942.1"
 /db_xref="GI:41053010"
 /translation="MEDWDSDFQPAVPSAKAPLKSKMADEVEDVDKSWEEER
 EKPRPPVEKTAAPKPSGAKKQQAQSTSEVVDALDDPALKEKRLQRLVEAD
 FKSSTTELEKDGKSGESKSLDTFIPKSESPFAVEALINKLPYKESFHYMGLKVMR
 LSMALSKGADKDISSTIAALANKEIKAKKEAAGKKQAKKQLHIENKDDPFIG
 KGNFDDPDPEVDFM"
 complement(10186..13754)
 /gene="P0575F10.2"
 complement(join(10186..10479,10759..10892,11127..11991,
 12089..12349,12515..12712,12843..13073,13455..13754))
 /gene="P0575F10.2"
 /note="start and end point are not identified"
 complement(join(10186..10479,10759..10892,11127..11991,
 12089..12349,12515..12712,12843..13073,13455..13754))
 /gene="P0575F10.2"
 /note="contains EST(s): AU058423(B51136), AU183658(B51136)"
 /codon_start=1
 /product="placitid division protein precursor"
 /protein_id="BAD07942.1"
 /db_xref="GI:41053011"
 /translation="MEGFHNLARPNSAPFASLPFRPRRRPPPHSAACRAAR
 WAERLFDLHLPRAESDPPSPAPAPASASPPVLPDPAERSLPLVDFFKVL
 GAEHFLGDDGIRRAFEARIAKPPQYGVSTDLVGRROWLQAHPTLANKNSRTQYDRA
 LNSRREBLTMDIMDKBEGALAVLVYVGGOLLDLPKRRKQVVLAMALAYDLIS
 DAMASPPDVGCCEVLERALKLQEDGASNLADLDLSQIDETLEITPRCVBELLS
 PIDEHKKRQEGQGANILMSVGRGSLAVVGGSFEAPMEAFELWSTIEQDF
 SKTPNSLPPEFELYNVALAHVAQALISKRQPTMMADLEQLQKFNIGSHVADNE
 MDLLEAFACSLIVGDSKCMWGLINSESSPPYDPKLIBETVYNSISISENDLPG
 CKLEETMLIFVPPSRDTRGMOPRLGYYDDPYVLSTLEMEGGASHLAAAAIAK
 LGAATATAITGVKSNATQAFMKVPLLEQDRLMENTKDGQVGLNPFDEMPAH
 SRNALKISKALPALAVYGATLPKRPKLSAIRSHGSVAVANSVSDSDPDLADE
 DPVHPNDADLADIVKWKOSIVSKALGPESHVASLOEVDGNLKWTRAEIER
 HGMEWETLSDVTIDISITISIDGRRAVEATIDEGQLDVTBPRNDSYDKTTRY
 EMASFSGKGMKTGAVLKS"
 complement(16401..17341)
 /gene="P0575F10.3"
 complement(join(16401..16739,16841..17341))
 /gene="P0575F10.3"
 /note="start and end point are not identified"
 complement(join(16401..16739,16841..17341))
 /gene="P0575F10.3"
 /note="similar to Arabidopsis thaliana chromosome 3,
 At3g61870"
 /codon_start=1
 /product="hypothetical protein"

CDS

mRNA

gene

gene	protein_id="BAD07943.1" /db_xref="GI:41053012" /translation="MMAAPLAAVHAITTCASNNKSPSPARQQOQTTTATRGSPPA LPSLDTAAAAATATLALAPDPDALLAGSEFGILLEGRSVALIHPVLMGGLPAILMAG YSGMORRVRTIODEINELKKQKPPAAAGATPAVAAGDSSSSSSPPSPAPKSVET KIDLETERRKKIKGSFRDRHFNAGSILLGCVTSVCGALNTWPRCKLPKPGHPLPA GAAITVLMAAAAALVPAMQKGNETARSILHAIINALINVLFTWQITFGLEIVKQPEFT TP"
gene	complement(18499, .18942) /gene="P0575F10.4" complement(18499, .18942) /gene="P0575F10.4" /note="start and end point are not identified" /gene="P0575F10.4" /note="contains EST(s) : AU184851(R10374), AU184850(R10374) " /codon_start=1 /product="putative EF-hand Ca2+-binding protein CDP1" /protein_id="BAD07944.1" /db_xref="GI:41053013" /translation="NMHQOIRSTTAAEQOASAGGGGEEYEDLMPWAGRLGAEGL LSLRAGFLPADPARGATIAELRSASVGLGGGEGEMTEBAAAVREGDQGD GALSEAFCVLMVRLSPGIMGDAEGWLEAIADLLRSLPPPPA"
mRNA	21937, 23552 /gene="P0575F10.5" join(<21937, .22225, 22326, .22648, 22736, .23049, 23193, .23552) /gene="P0575F10.5" /note="start and end point are not identified" join(21937, .22225, 22326, .22648, 22736, .23049, 23393, .23552) /gene="P0575F10.5" /note="similar to Arabidopsis thaliana chromosome 5, At5g24330"
CDS	/codon_start=1 /product="hypothetical protein" /protein_id="BAD07945.1" /db_xref="GI:41053014" /translation="MGPARPLRRRTARRPAAATRAEGSGGGDDDDVCEACSGGESA ABELLDCDGRGLHIFCLRPILPRVAGWPCPCGASPSPHSKSHAAKKEPFLVQ TKIVDFKIORGPAAALAAASSEGGKKRRVGGRIYVSKKKRLCLFNSDDPAR LRMAATLALATGATGAVFNSNLTIVPGMAPRANAALAESGGMOVLPREDEVETINICK RMARBERPPLVYDPRVGFTVEADRFKIDLTITTEYGVGDVDTLRHHDGSMAT LISLAATPSRLVTCPDKRSNIARFTINGNHTPDRCKONLKCVRFDVGSGCRVLVA NRLISGRLVYDSEHEPHTHFV"
gene	complement(10124479, .24817, 24913, .25266, 25398, .25633, 25720, .25832, 27054, .27185, 27995, .28147) /gene="P0575F10.6-1" complement(10124479, .24817, 24913, .25266, 25398, .25633, 25720, .25832, 27054, .27185, 27995, .28147)) /gene="P0575F10.6-1" /note="supported by full-length cDNA(s) : AK121372" complement(10124491, .24817, 24913, .25266, 25398, .25633, 25720, .25832, 27054, .27185, 27995, .28147)) /gene="P0575F10.6-1" /note="supported by full-length cDNA(s) : AK062255" complement(24563, .28147) /gene="P0575F10.6-2" complement(24563, .28147) /gene="P0575F10.6-2" /note="contains EST(s) : AU031017(BE0571), AU093466(BE0571) contains full-length cDNA(s) : AK121372, AK062255" /codon_start=1 /product="putative arginine/serine-rich aplicing factor
mRNA	25720, .25832, 27054, .27185, 27995, .28147)) /gene="P0575F10.6-2" complement(24563, .28147) /gene="P0575F10.6-2" /note="contains EST(s) : AU031017(BE0571), AU093466(BE0571) contains full-length cDNA(s) : AK121372, AK062255" /codon_start=1 /product="putative arginine/serine-rich aplicing factor
misc_feature	non-coding transcript probably inactive due to including stop codon(s) in CDS" complement(10124787, .24817, 24913, .25266, 25398, .25633, 25720, .25832, 27054, .27185, 27995, .28147)) /gene="P0575F10.6-1" /note="contains EST(s) : AU031017(BE0571), AU093466(BE0571) contains full-length cDNA(s) : AK121372, AK062255" /codon_start=1 /product="putative arginine/serine-rich aplicing factor
gene	complement(24563, .28147) /gene="P0575F10.6-2" complement(24563, .28147) /gene="P0575F10.6-2" /note="contains EST(s) : AU031017(BE0571), AU093466(BE0571) contains full-length cDNA(s) : AK121372, AK062255" /codon_start=1 /product="putative arginine/serine-rich aplicing factor
CDS	8.7%; Score 317.6; DB 15; Length 150462; Best Local Similarity 49.3%; Pred. No. 1,5e-60; Matches 1459; Conservative 0; Mismatched 1299; Indels 202; Gaps 16;

OY	744	TGATTTCACAGGTATTTAGAGCTCAAAACAATTTCTTAACTCGATTCGAATCAGAAGGC	803
Db	13086	TGGCTTTTGCTCAGGTTCTTAGGGGCGAAGGCCAACATTTCTTGCCGATGCGATCAGAAGGGC	130277
OY	804	ATTTCGAAGCTAAGGTTTTCGAAAACCGCGGAAATTCGGTTTTCAAGGACGACCGCTTTATCAG	863
Db	13026	GTTTCGAGGCAACGATAGCCAGACGACCGCATGTGGCTACACACGAGATGCTCTTGTTGG	12967
OY	864	CCGAGACAGATTTCTTCAAGCTCTTGCGAAACTCTGTCTAACTCTCGGCTTCAAGAGA	923
Db	12966	TTCGCGACAAATCTCGACAGATTCGCCATGACACTCTCATGAACCAGAACTCCCGCATCA	129070
OY	924	GTCACAAAGAGCTTCTTGATATGATAAGATCACAGTACATCATGATGTTCTTGCGGA	983
Db	12906	GTAATGATCGTCCCTTTCTGAGAACCGTAAAGAGCTTCACCATGATATATGCTTGGGA	12847
OY	984	TAAAGTAATTCGATTTCCGAAATATAAAGTTCT-----TCGT	1022
Db	12846	CAGGTGAGGCTTTAAGGGGCGACCAAAGGTGTGTGATCTGTTCAAATAGCACTAGT	127877
OY	1023	TTTATATTCATGAAATTTGATTAAGGAAGACCTTTATCTAGTGAAGTTCTCGGGGCTC	1082
Db	12786	GTAATACTCTCTCTGCTTATATATATATTTGGGGTCTCTGTGAAGAGGTTCCAGGTGTGT	127277
OY	1083	TCGTGTATTTGCAAGAAGGTGTGAGACTAGATAGTTCTTCGGGTGTGAGGCTCTGC	1142
Db	12726	TGTGTGGGCTTCAGAGAGCTCGGGGAAGCACTGTGTGTTCTTMACTGSAAGAACGTTGC	12667
OY	1143	TTAAGAGAGGTTGCCCTTAAGTCTTTAAGCAAGATGTGGTTTTAGTTATGCGCTTGCGT	1202
Db	12666	TTCTGATTCGAGCCACCAACGAGCTTCAACACAGGACCTGTGTCTAGCGATGCGCTCGGCTT	126070
OY	1203	TTCTCGATGTCTGAGGGAATGCTATGGCAATGGATCACTGTATTTTATTACTGTGTATNG	1262
Db	12606	ATGTGATCTTATCAGAGGAATGCTATGACAGCAAGCCCTCAGATGTAATTTGGCTCTGCG	12547
OY	1263	AGTTTGTGAAGAAAGCTTGAACCTTTTACAG-----TAG	1298
Db	12546	AGGTGCTCGAAGAGGCTCTCAAGCTCTTGCAAGTGAACGTTTGCTTTTACTCCCTAC	12487
OY	1299	TTTGCATCTGCTTTGGTATTTTGCACAGCGTGTGCTTATAPAGAACTTCTTGATTTG---	1355
Db	12486	TYTTTCATTTGGGCTTATATTTTGAATGGAATPAGATTAATAAACCTGTGAMAATTTGGTTT	12427
OY	1356	-----ATACCTTGTATTTGAGTCTTG-----	1376
Db	12426	GGGACATGAGGCTATATTTTTTTATATAGAAATGGTGAATTCCTTGAAACATATTAATA	12367
OY	1377	-----TGTMGAGGAAGAGCAATGACCTTGACCCGATTTTACGTGCACAA	1424
Db	12366	GGAATTCATATCTGTGGAAGATGAGGACGAAAGCAATCTCCACACTGATCTGCTTTCACAGA	123070
OY	1425	TTGATGTGATCTTTGGAAGAGATACATCCGCGTTATGTCTTGGAGCTATCTGGCTTACCGC	1484
Db	12306	TTGATGTGAATCTCTCGAGAGATTKACCTCGCTGTATTTGAGAGCTTCTCTCCCTCTA	12247
OY	1485	TTTGATGTGATTAACGCTGCGAAAAAGCTAAATGTGTTTAAAGCGGTGTGCGGAATATTTGT	1544
Db	12246	TTTGACACAGAGCAATCTAATGAAGACCGCCAAGAAAGGCTTCAAGSTGCGAAMAACTTTTGT	121877
OY	1545	GGTCTGTTTGAAGAGATGAGCATCAAGCTCTTTGGGGGCTTGAACCGTGAAGATTTTA	1604
Db	12186	GGAGCTTTGCAAGAGAGGATTTGCTTACCGTTTGAAGAGAAATTTTCTGTGAAGCTTCA	121277
OY	1605	TGATGTAGGCGTTTTTACGATGACAGCTGTGAGCAGGT-----ATACAGTTTAA	1655
Db	12126	TGAACGAGGCTTTTGGAGATGACATCAATTTGAACAGGTTGTGTAACATGTTGTTTTAT	12067
OY	1656	AATACCTTTTATATTTCTTTAGCATGATATACTTTAGGTTTCTCATTTTAAATGATGT	1715
Db	12066	GAGCAACATCTCTATATTTATTTGACAGAGACACATCTTTGAAAAACATGTTAATCTCCCTTC	120077

QY 1716 TGT-----GTGGAGTGTATCTTTTGTAGTACCCCAAGCAATATCCAGAGTCA 1770
DB 12006 TGTACACCTCCAGATGATTTCTTTCAAAAACCGAATAGCATTTCTCTCGATGG 11947
QY 1771 TTGAAAGTTAGAAATTGCACTTGTCTTGTGGCTCAAGCTTTTATGTGAAGACCA 1830
DB 11946 TTGAAATTTACAAATGACACTGTGCAATGTGCTCAAGCAATTAATGAAGGCCA 11887
QY 1831 CACCTTTACAGAGTCTGATTAAGCAATTCAGCACTTCAGAGGCTAAGGTAATGCT 1890
DB 11886 CATTCATCATGATGCGGATGATCTTTTGAACAACCTCAAGAGTTCAATAGTTCT 11827
QY 1891 ATGAGATCTCTGCGATGTTGATGATACCGAATTAATGGAGATTAAGCTTGGCTA 1950
DB 11826 CATTAATGCTTATG-----ATATAGAGATGACCTTGATG 11791
QY 1951 GAAAGGGGCTCTGTCATCTTTAAGCAAAAGTTGATGATGCGGTATGTGGTGGC 2010
DB 11790 GAAAGGGCATCTGCTCATTTGCTAGTGGAGATGTTAGCAAGTGCAGAAATGGCTTGA 11731
QY 2011 TTAGACAGTGAAGATTCACATATAGAAATCCAGCTATGTGAGTTGTTTGGAGAT 2070
DB 11730 ATTGATTAATGAGCTTCCACATACAGAGACCCCAAAATTTAGAGTTATGTGACCAAC 11671
QY 2071 TC---AAATGTGATGACAAATGATGATCTCTGACATAGCAAAATTTGTTGAACCTGG 2127
DB 11670 TCTAGCATGATGAGAGATGATCTTCCAGGGGCTGTGCAAGCTTTTGGAGACTGG 11611
QY 2128 TTGGCAGGGGTGTCTTCTTCTAGTTCAAGACACCAAGATTAATAAATTTAACTGGG 2187
DB 11610 CTATCTTTGAGGTTTCTTGGAGACAGATATCTGGGGCATGCAAGTTGCACTTGA 11551
QY 2188 GACTACTATGATGATCTTATGTTTGAATCTTCTTGAAGAAGTGGAGGTAGTTCAGGT 2247
DB 11550 GATTACTAGAGATGATCAGAAATTTTAACTTACCTAGAGAGATGAGGGGTGTGTCT 11491
QY 2248 TCTCTTATGCTGTCTGCACTATGCAAGATTTGAGC-----C 2289
DB 11490 TCTCATTTGGTGTCTGTCTGTATGCAAACTTGGTGTCAAGCTACAGTGCATCT 11431
QY 2290 GAGCATGTGAAGCTAGTGTCTATGACAGGACCTGCAAGAAATTTTCTTCCGCTATAC 2349
DB 11430 GGTACTGTGAATCAAAATGCTATGCAAGGTTCAACAGGTTTTCATTTGATAGAACG 11371
QY 2350 GATAGAAATCTGGCTGAACCCAGAGATGTCAAGACAGTGTATAGTGTAGATCTGT 2409
DB 11370 TTAGACAGGTGACCCATGAAAAATCTAAAGATGCGCTTGGGGATATTTGAAAAATTT 11311
QY 2410 GGTAAATGATGAGCGGTGATGATGCTGTG-----TGTCTTAT 2451
DB 11310 GACCAAGAAATGCACTGTCTATGATTTGAAAAATGCGCGCTTGAAGATTAATCTGTCT 11251
QY 2452 GCAGAACTGTAAAGCCTCTGAAAACTTTGAACCTATGATTAATGCAATTCGAGCTGG 2511
DB 11250 GCGGCACTGTGTGCACTGTGTGCAAGTAATTTGGGCAATATTTGCTGTGAAGGCC 11191
QY 2512 GCTCAGAGAGAGGCTGATGAACTACTGTTGAATGCGGTGCTGATATGTTAAAG 2571
DB 11190 CTTTCTGCTATTAAGAGTACAGATGATTTGTGCGAGTTGCTAATTAATGTGCACTTACT 11131
QY 2572 GAGGCAAGTGAAGATCTGAGTGTGTGAGCAATTTGAATTTCACTGTTTCAAGC 2631
DB 11130 GATGTGATTAATGCTTTTCAAT--TGCTAAATGATGCAACACATGGAATTTACTG 11073
QY 2632 CAGAAATATTTCTTAAAGCAGCTCATCTTTCAACGCAAGATATGTTCTTCTATG 2691
DB 11072 CTCACGTGTTTACGCGTATCTTGTGCTATTAAGATATGTTCTTATCTTATGCT 11013
QY 2692 GAATCTGATGTGCTACCATATGATTAATGATTAATGCAATTTTCAATATCTGATG 2751
DB 11012 TTGCATTAATGTTTAAACAACTATACATAATTTTGTCTCATACATACAAATG 10953
QY 2752 CTCAAAAATATGCTTGTGTTGTGAGCTAAGACATGTTCCACTTAATATACATGTCCAAA 2811

DB 10952 CTACTATTAATCTGTAAATGATATCATGAGAGTTAACTTTTGCATGATGATGTTGAAG 10893
QY 2812 AGTTTACCAAGATTAACAGATGCTGATTAATTTCACTAATATGCTGTG-----A 2866
DB 10892 ATCTGCACTGATATAGATTCAGATATTTCTTAAGATGATGCAAGCTGGCAGAAAG 10833
QY 2867 ATTTTGTATCAAACTGTAGACAGAAATGTAATTTCACTCTCAATTTCTGTTTAAAG 2926
DB 10832 ATATTGTTGCAAGGCGAGATATCAATCTTAAGCCCTGGGACCAAGAACTTGGTTG 10773
QY 2927 TAAGTATGATTAAGATTTGCTTATGTTGCTTGTGCTTGTCCAACTTTCTTTT-CTTGATTT 2985
DB 10772 CATCATTCAGAGAGTGTATGTTTAAACTTTGACTTTTCACTGCTTCTAGCTGGGCACCTG 10713
QY 2986 TTTTCTTTTCAATTAAGGCTAGTGCAGAGCTGACAGATCAGAGCACTCCAGAAATGA 3045
DB 10712 CTCTTACTATTAATCTTACCTTCTTATAGAGATTAATTAAGATTAATCTTCTTATAGGA 10653
QY 3046 TGTAGACCTGACAGAAATATATATCCAAAGTGCAGAAATTAAGTCTGTGGCTTTTGG 3105
DB 10652 TATATGTCACAGCAATGATACCTTTACAAATTTGTTTCAATTAATCTTTCT----- 10602
QY 3106 GCTGATACCGCATAGAAATGTTACCAAGAGTGAAGGAATTAATCTACATTCATCA 3165
DB 10601 ---TGTGTACTCACACATGCTTCAATTCGAATGCAATTTTCAAGAACTATTTGC 10545
QY 3166 TTGTGTAAACCTGTGTGACATGATTAATGATGCTGTGTGCTTGTGATTT---CTGTATTT 3222
DB 10544 CACTGCATGATATATCCCTGTATGATGCGGTCAAAATGTGTACTGATTTTCCATGTTCT 10485
QY 3223 TATAGTTTGTGATGAGGCAATGCTGAAGATTTTGAATGCAAGACAGCACTGCAAGCTGGC 3282
DB 10484 TATAGTTCTTGTATGCAACATGCTTAAGGTGTGATGCAAGCAGACAGGAGATTGAGC 10425
QY 3283 AGCTTGGTGTGTTATATATTAATACATCTTTGAACATATCTGTTGACAGTGTGACATCT 3342
DB 10424 GTCAATGGGTGTGTTGCGAGATATACATATCCAGTGTACATGATTAATGATCACTATCT 10365
QY 3343 CAGCAGATGAAACCGGTCTGTGTGGAAGCACTGTGAGAGCTGTGTCTATCATCTG 3402
DB 10364 CCTAGATGTGTGACAGACAGCTGTGAGGCTTACATTAATGAGGAGGCAACTTACTG 10305
QY 3403 ATTTGTTTATTCAGAAAAATGCTACTGATGTGCAAGAACTTACCAACAAGATTCGAAG 3462
DB 10304 ATGTTACTGAGCCCAAGAAATGATTCATATGACAAATTAATCACTCCGATATGGA 10245
QY 3463 TTTTGTGTCAAG---TGAGGTTGAATAATCACTGAAGGCTGTGTTTGTGATCATAT 3519
DB 10244 TGGCTTCTCCAAAGCTAGAGGGTGAAGATTAACGAAGAGACAGTCTCAAGTGTAGA 10185
QY 3520 ATACTCATATGATGATGTC 3539
DB 10184 ACAGTGTGCAAGTACCTTGTG 10165

RESULT 10
AC155597
LOCUS 184752 bp DNA linear HTG 25-JAN-2005
DEFINITION Zea mays strain B73 clone ZMMB0216K08, *** SEQUENCING IN PROGRESS
ACCESSION AC155597
VERSION AC155597.2 GI:58082456
KEYWORDS HTG; HTGS_PHASE1.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
1 (bases 1 to 184752)
Chan, A.P., Pertea, G., Zheng, L., Cheung, F., Lee, D., Koo, H.,
Utecherback, T.R., Feldblum, T.V., Rabinowicz, P., Fraser, C.M.,
Schubert, K., Samtignel, P., Ma, J., Pontaroli, A.C., Rohlfing, T.,

TITLE
JOURNAL
REFERENCE
AUTHORS
Budiman,A., Bedell,J.A., Lakey,N., Barbazuk,B., Bennetzen,J. and Quackenbush,J.
Consortium for Maize Genomics - BAC skim sequencing and assembly
Unpublished
2 (bases 1 to 184752)
Chan,A.P., Pertea,G., Zheng,L., Cheung,F., Lee,D., Koo,H., Utecherback,T.R., Feldblum,T.V., Rabinowicz,P., Fraser,C.M., Schubert,K., Samiiguell,P., Ma,J., Pontaroli,A.C., Rohlfing,T., Budiman,A., Bedell,J.A., Lakey,N., Barbazuk,B., Bennetzen,J. and Quackenbush,J.
Direct Submission
Submitted (15-JAN-2005) The Institute for Genomic Research (TIGR), 9712 Medical Center Dr, Rockville, MD 20850
3 (bases 1 to 184752)
Chan,A.P., Pertea,G., Zheng,L., Cheung,F., Lee,D., Koo,H., Utecherback,T.R., Feldblum,T.V., Rabinowicz,P., Fraser,C.M., Schubert,K., Samiiguell,P., Ma,J., Pontaroli,A.C., Rohlfing,T., Budiman,A., Bedell,J.A., Lakey,N., Barbazuk,B., Bennetzen,J. and Quackenbush,J.
Direct Submission
Submitted (25-JAN-2005) The Institute for Genomic Research (TIGR), 9712 Medical Center Dr, Rockville, MD 20850
On Jan 25, 2005 this sequence version replaced gi:57863118.
----- Trace Submission -----
Center name: TIGR
Seq lib id: ZGGO
----- Project information -----
Web site: <http://www.tigr.org/cdb/tgi/maize/>
Contact: maize@tigr.org

NOTE: This is a 'working draft' sequence. It currently consists of 26 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
1
2210: contig of 2210 bp in length
* 1
* 2211 2310: gap of unknown length
* 9942 9941: contig of 7631 bp in length
* 10042 10041: gap of unknown length
* 16745 16744: contig of 6704 bp in length
* 16845 16844: gap of unknown length
* 32394 32393: contig of 15549 bp in length
* 32494 32493: gap of unknown length
* 32495 32494: contig of 3000 bp in length
* 35594 35593: gap of unknown length
* 35595 35594: contig of 4457 bp in length
* 40051 40050: gap of unknown length
* 40151 40150: gap of unknown length
* 41379 41378: contig of 1228 bp in length
* 41479 41478: gap of unknown length
* 44517 44516: contig of 3038 bp in length
* 44617 44616: gap of unknown length
* 44618 44617: gap of unknown length
* 45564 45563: contig of 947 bp in length
* 45665 45664: gap of unknown length
* 49440 49439: contig of 3776 bp in length
* 49540 49539: gap of unknown length
* 52698 52697: contig of 3158 bp in length
* 52798 52797: gap of unknown length
* 55964 55963: contig of 316 bp in length
* 55965 55964: gap of unknown length
* 59236 59235: contig of 3172 bp in length
* 59336 59335: gap of unknown length
* 59337 59336: contig of 10716 bp in length
* 70052 70051: gap of unknown length
* 70053 70052: gap of unknown length
* 84603 84602: contig of 1451 bp in length
* 84703 84702: gap of unknown length
* 86418 86417: contig of 1715 bp in length
* 86518 86517: gap of unknown length
* 88069 88068: contig of 1551 bp in length
* 88169 88168: gap of unknown length
* 89661 89660: contig of 1492 bp in length

FEATURES
source
* 89662 89761: gap of unknown length
* 89762 91357: contig of 1596 bp in length
* 91358 91457: gap of unknown length
* 91458 132727: contig of 41270 bp in length
* 132728 132827: gap of unknown length
* 132828 137468: contig of 4641 bp in length
* 137469 137568: gap of unknown length
* 137569 163843: contig of 26275 bp in length
* 163844 163943: gap of unknown length
* 163944 169858: contig of 5915 bp in length
* 169859 169958: gap of unknown length
* 169959 176100: contig of 6142 bp in length
* 176101 176200: gap of unknown length
* 176201 183079: contig of 6879 bp in length
* 183080 183179: gap of unknown length
* 183180 184752: contig of 1573 bp in length.
Location/Qualifiers
1. .184752
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/c1one="ZM58B0216K08"
2211. .2310
/estimated_length=unknown
9942. .10041
/estimated_length=unknown
16745. .16845
/estimated_length=unknown
32394. .32494
/estimated_length=unknown
35495. .35594
/estimated_length=unknown
40052. .40151
/estimated_length=unknown
41380. .41479
/estimated_length=unknown
44518. .44617
/estimated_length=unknown
45565. .45664
/estimated_length=unknown
49441. .49540
/estimated_length=unknown
52699. .52798
/estimated_length=unknown
55965. .56064
/estimated_length=unknown
59237. .59336
/estimated_length=unknown
70053. .70152
/estimated_length=unknown
84604. .84703
/estimated_length=unknown
86419. .86518
/estimated_length=unknown
88070. .88169
/estimated_length=unknown
89662. .89761
/estimated_length=unknown
91358. .91457
/estimated_length=unknown
132728. .132827
/estimated_length=unknown
137469. .137568
/estimated_length=unknown
163844. .163943
/estimated_length=unknown
169859. .169958
/estimated_length=unknown
176101. .176200
/estimated_length=unknown
183080. .183179
/estimated_length=unknown

ORIGIN

Query Match 6.7%; Score 246.8; DB 14; Length 184752;
 Best Local Similarity 57.0%; Pred. No. 1.3e-44;
 Matches 564; Conservative 0; Mismatches 372; Indels 54; Gaps 4;

QY 1329 TGGCTTATAGAACTCTTGTGATTTG--ATACTTGTATTGAGTCTTGTGAGAGGA 1386
 |||||
 DB 108883 TGTCTTGATTAAGATATCTTGTAAGCACTACTAATGAAACATTTTCTTTAGAGGA 108942

QY 1387 AAGAGAACTAGCCTTGACCCGATTTAGTGCACAAATTTGATGAGCTTTGGAAGAT 1446
 |||||
 DB 108943 TGGTGAACAAATCTTGACCTGAACTGTTGACAGATTGAACTTTGAGAGAT 109002

QY 1447 CACTCGCCTTATGCTTGTGAGCTACTTGGCTTACCGCTTGTGATGATTAGCTGCNA 1506
 |||||
 DB 109003 TACACTCGTGTGATTTGAGACTTCTGCTCTCTACTGATGAAAAATTA 109062

QY 1507 AAGACTAAATGCTTTAAGCGGTGCGGAATATTTTGTGCTGTTGAGAGAGTGGAGC 1566
 |||||
 DB 109063 ACGCAAGAAAGCTTGCAAGGTGCAAAAAACATATTTGAGATGTTGGCAGGGGTGTAT 109122

QY 1567 ATCAGCTCTTGTGGGGGTTTGAACCGTGAAGAAATTGATGAGGCGTTTATAGAT 1626
 |||||
 DB 109123 TGTCTACTGTTGAGAGAGATTTTCTGTGAGGCTTCATGAATGAGGCTTCTTGCAGAT 109182

QY 1627 GACAGCTGCTGAGCAGGT-----ATACAGTTTATGATTA 1661
 |||||
 DB 109183 GACATCAGCTGAGCAGGTTTGTGTCTACCAAGTTTAATGACAACTACTGTATGTGTA 109242

QY 1662 TTTTATTTCTTTAGCAATGATTAATTAAGTTTCTCATTTTAATGATGTTGTG 1721
 |||||
 DB 109243 GCTCTTCCATTTTACTTGAACACAGTTTATAGAGCTTCACTCATTTTCTGTATATC 109302

QY 1722 GTAGGTGATCTTTTGTAGCTAACCCCAAGCAATATCCAGAGAGTCACTTGAAGTTA 1781
 |||||
 DB 109303 TCAGATGATTTCTTCTTAAACACGATATGCAATCACCTGATGATGTTGAATCTA 109362

QY 1782 CGAATGTCAGCTGCTTGTGTGCTCAAGCTTTTATTTGTAAGAACCCACCTTTTACA 1841
 |||||
 DB 109363 TAGTGTGCACTTGTGCCACATTTGCTCAAGCAATTCGAATTAAGGCCCAATTCATAT 109422

QY 1842 GATGCTGATTAAGCAATTCAGCAACTTCAGCAGCGCTAAGTATGCTATGAGAAATTC 1901
 |||||
 DB 109423 GATGCAAGTGAATCTTCTGAAACAATTGAGAAATGATAGGTTCTCAATATCTCTTA 109482

QY 1902 TGGCATGTTGATGATCAACGGAATATTTGGAGATAGACTTCGCTTGAAGAAAGGAGCT 1961
 |||||
 DB 109483 TGAATAAC-----GAGATGACCTTCGATTAAGAAAGGAGCT 109518

QY 1962 CTGTGCACTGCTTATAGCAAGTATGATGATGCGTATGCTTGGGCTTGAAGAGTGA 2021
 |||||
 DB 109519 CTGCTGCTTACTTGTGGGGGATATTTGCAATTTGCAATGTGCTTGAATTTGATATGA 109578

QY 2022 GATTTCAATATAGGAATTCAGCTATTTGTGAGTTTGTGGAATTC---AAATCG 2078
 |||||
 DB 109579 GTCTCTGCTATATAGGAGACCCCAAAATATCGAGTTTGTGATTAATCTCAAGCATGA 109638

QY 2079 TGATGCAATGATGATCTCCCTGACATATGCAAAATTTGGAACCTGTTGGCAGGGGT 2138
 |||||
 DB 109639 CGAAAGATATGATCTTCTTCCAGGGCTGTGCAAGCTTGTGAGACATGTCTGTCCGA 109698

QY 2139 TGTCTTCTTAGGTTCAAGACACCAAGATTAATACTCGGGGCTACTATGA 2198
 |||||
 DB 109699 GGTATTTCCAGAGAGACAGATATCCGAGGAGCAGATTTGAGCTGGGAGACTATTTAGCA 109758

QY 2199 TGATCTATGTTTGTGAGTTTCTTGAAAAGATGAGGTGATTCAGGGTTCTCTTTAGC 2258
 |||||
 DB 109759 CGACCCAAAGTTCTTAGCTATAGAAAGATGAGAGCGGTGTGCTTCCATTTAGC 109818

QY 2259 TGTGCTGCACTATGAGCAAGATTTGGAGC 2288
 |||||
 DB 109819 TGCAGCTGCGCTATAGCAAACTTGTGTGC 109848

RESULT 11

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

AUTHORS

TITLE

JOURNAL

D0022571 550 bp DNA linear STS 28-MAY-2005
 Beta vulgaris chromosome 1, sequence tagged site.

D0022571.1 GI:66394762
 STS.

Beta vulgaris
 Beta vulgaris

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
 Caryophyllales; Amaranthaceae; Beta.

1 (bases 1 to 550)
 Friesen,T.L., Weiland,J.J., Ashheim,M.L., Hunger,S., Borchardt,D.C.
 and Lewellen,R.T.

Identification of a SCAR marker associated with Bm the Beet mosaic
 virus resistance gene on chromosome 1

Unpublished
 2 (bases 1 to 550)
 Weiland,J.J. and Friesen,T.L.

Submitted (29-APR-2005) Sugarbeet and Potato Research,
 USDA-Agricultural Research Service, 1307 18th St. N, Fargo, ND
 58105, USA

Location/Qualifiers
 1..550
 /organism="Beta vulgaris"
 /mol_type="genomic DNA"
 /cultivar="C719"
 /db_xref="taxon:161934"
 /chromosome="1"

/note="fw primer seq: GACTGAGTGGTGAAGCAAGTGTG;
 rev primer seq: GGAAGCAATTCATCTTTATAGT;
 fwd_primer_name: Rbm05fwd; rev_primer_name: Rbm05rev"
 1..550
 /note="SCAR marker associated with resistance to Beet
 mosaic virus"

STS

ORIGIN

Query Match 4.6%; Score 170.4; DB 10; Length 550;
 Best Local Similarity 75.0%; Pred. No. 2.7e-27;
 Matches 213; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

QY 1361 TTGTTATGAGCTCTGTGTAGAGAGAGCAAGTATGCTTGAACCGGATTTACGTCA 1420
 |||||
 DB 96 TTCTGATCTACTATTTGTATAGAGAGGCTTCAAGTATGCTTGTGCGCGGATTTACAGCA 155

QY 1421 CAAATTTGAGACTTTGGAAGATGATCCTCCGGTTATGCTTTGAGACTTGGCTTA 1480
 |||||
 DB 156 GAGGTGACGAGCATTTGAGAGATGATGAGATGATGATGATGATGATGATGATGATGATGAT 215

QY 1481 CCGCTGTGATGATTTACGCTGCCGAAAAGACTAAATGTTTAAGCGGTGTGCGGAATTT 1540
 |||||
 DB 216 CCTCTCAGTATGATTTTTCGGAAGAAAAGATGAAAGGCTTCAACGGGTATGCGAATATA 275

QY 1541 TTGTGCTGTGTGAGAGAGGTGAGCATCAGCTCTGTGTGAGGGTTTGAACCGTGAAG 1600
 |||||
 DB 276 TTATGCTGTGTGAGAGAGGTGAGCGGACGCTATAGCTGTGTGTGTTTACTCTGTGAGAT 335

QY 1601 TTATGATGAGAGCGTTTATGAGATGACAGCTCTGAGCAGGT 1644
 |||||
 DB 336 TTCAATGATGAGGCTTTTTCATATGACAGCAGCTGAGCAGGT 379

RESULT 12
 AC160013
 LOCUS 117818 bp DNA linear HTG 02-JUN-2005
 DEFINITION Medicago truncatula clone mtn2-162b23, WORKING DRAFT SEQUENCE, 4
 ordered pieces.
 ACCESSION AC160013
 VERSION AC160013.10 GI:66865008

KEYWORDS	HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE	Medicago truncatula (barrel medic)
ORGANISM	Medicago truncatula
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eustosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae; Medicago.
TITLE	1 (bases 1 to 117818)
JOURNAL	Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B., Cook,D., Kim,D. and Roe,B.A.
AUTHORS	Medicago truncatula BAC Clone mch2-162b23
REFERENCE	2 (bases 1 to 117818)
JOURNAL	Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B., Cook,D., Kim,D. and Roe,B.A.
AUTHORS	Direct Submission
REFERENCE	Submitted (16-APR-2005) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA
TITLE	3 (bases 1 to 117818)
JOURNAL	Lin,S., Dixon,R., May,G., Sumner,L., Gonzales,B., Cook,D., Kim,D. and Roe,B.A.
AUTHORS	Direct Submission
REFERENCE	Submitted (02-JUN-2005) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA
COMMENT	On Jun 2, 2005 this sequence version replaced gi:66841540.
TITLE	----- Genome Center
JOURNAL	Center: Department Of Chemistry And Biochemistry
AUTHORS	The University Of Oklahoma
COMMENT	Center code:UOKNOR
FEATURES	-----
SOURCE	* NOTE: This is a 'working draft' sequence. It currently
FEATURES	* consists of 4 contigs. Gaps between the contigs
FEATURES	* are represented as runs of N. The order of the pieces
FEATURES	* is believed to be correct as given, however the sizes
FEATURES	* of the gaps between them are based on estimates that have
FEATURES	* provided by the submitter.
FEATURES	* This sequence will be replaced
FEATURES	* by the finished sequence as soon as it is available and
FEATURES	* the accession number will be preserved.
FEATURES	1 18815: contig of 18815 bp in length
FEATURES	* 18816 18915: gap of unknown length
FEATURES	* 18916 21097: contig of 2182 bp in length
FEATURES	* 21098 21197: gap of unknown length
FEATURES	* 21198 35201: contig of 14004 bp in length
FEATURES	* 35202 35301: gap of unknown length
FEATURES	* 35302 117818: contig of 82517 bp in length.
FEATURES	Location/Qualifiers
FEATURES	1..117818
FEATURES	/organism="Medicago truncatula"
FEATURES	/mol_type="Genomic DNA"
FEATURES	/db_xref="taxon:3880"
FEATURES	/clone="mch2-162b23"
FEATURES	/clone_lib="Medicago truncatula BAC library H2"
FEATURES	18816..18915
FEATURES	/estimated_length=unknown
FEATURES	21098..21197
FEATURES	/estimated_length=unknown
FEATURES	35202..35301
FEATURES	/estimated_length=unknown
ORIGIN	
Query Match	4.0%; Score 146.2; DB 14; Length 117818;
Best Local Similarity	68.5%; Pred. No. 6; 2e-22;
Matches 202; Conservative	0; Mismatches 99; Indels 0; Gaps 0;
3225	TAGCTTTTGATGGCGCAATGCTGAAGATTGGACTGACAGAGCAGCTGAATCGCGAG 32844
1084	TAGGTGTTGACGCGCGAATGTTGAAGATTGAGACTGATCGACGCGATTTGACAG 11433
3285	CTTGGGTGATTATGATTAACACTGTGGAACATATCTGTTGACAGCTGACAGTCTCA 33444

Db	1144	CTTGTTGGTCAATGACTACAACTTGGAAGATCTCAACATCGACAGTGGACATATCA	1203
Qy	3345	GCAATGGAACCCGTGCTCTGTGTGGAAGCAACTTGGAGAAGTCTGTTGTATCTGAT	3404
Db	1204	CAGATGGGGCGGCTTGGAGTAGTGGAAACACTTCAAGAGTCTTACCCACTTCAGTCT	1263
Qy	3405	TTGGTATCCAGAAACAATGCTACTGATGTGCAACTACACAAAGATACGAAGTT	3464
Db	1264	GTTGGTATCCAGCACTGCTACTTCCACAGCGAACTTACACAAAGATATGAATG	1323
Qy	3465	TTCTGTCTCAAGTCAGGGTGGAAAAATCACTGAAGGCTCTGTTCTTGATCATAT	3519
Db	1324	TCTTTTCAAGTTCAAGGGTGGAAAAATTTATGAAGAGACTGCTTGATGCTGTAAT	1378
RESULT 13			
LOCUS	166494	Sequence 14 from patent US 5670367.	linear PAT 28-DEC-1997
DEFINITION	166494		
ACCESSION	166494		
VERSION	166494.1	GI:2724471	
KEYWORDS	.		
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 7218)		
AUTHORS	Dorner, F., Scheifflinger, F. and Falkner, F. Gunter.		
TITLE	Recombinant fowlpox virus		
JOURNAL	Patent: US 5670367-A 14 23-SBP-1997;		
FEATURES	Location/Qualifiers		
SOURCE	1..7218		
	/organism="unknown"		
	/mol_type="unassigned DNA"		
ORIGIN			
Query Match	1.6%; Score 57.6; DB 6; Length 7218;		
Best Local Similarity	3.9%; Pred. No. 0.064; Mismatches 147; Indels 0; Gaps 0;		
Matches	15; Conservative 218; Mismatches 147; Indels 0; Gaps 0;		
Qy	510	TCTCTCCCATTCMAATTATGCGATTACACACGCGCAGCAAAAGCTCCGACTAGCCA	569
Db	1078	YY	1137
Qy	570	CAACACCTTACACATATCTGCTCCGACGAAATGGCCGACGCTTCTCTCCGACTT	629
Db	1138	YYY	1197
Qy	630	CAATTTACCGCATTCCTCCTCCTCCTGTCGACACGCGACACACCGCACCT	689
Db	1198	YYY	1257
Qy	690	CGTCTCTGCGACACATCTATGATGCTGCCGACGCGACGTCGCCCATTCGATGATT	749
Db	1258	YYY	1317
Qy	750	CTACACAGTATAGAGACTCAACACATTTCTTACCGATGGAATCAGAAAGACATTGCA	809
Db	1318	YYY	1377
Qy	810	AGCTAGGGTTTCGAAACGCGCGCAATCGGTTGAGGACGAGCGCTTATACGCGGAG	869
Db	1378	YYY	1437
Qy	870	ACAGATTTCTTCAAGCTGCTT 889	
Db	1438	CCAATTTCTTCTATCTCTTT 1457	
RESULT 14			
LOCUS	ARS79680	1141 bp DNA	linear PAT 14-DEC-2004
DEFINITION	Sequence 22 from patent US 6784342.		
ACCESSION	ARS79680		

VERSION AR579680.1 GI:56583130
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1141)
AUTHORS Kunst, L. and Clemens, S.
TITLE Regulation of embryonic transcription in plants
JOURNAL Patent: US 6784342-A 22 31-AUG-2004;
The University of British Columbia; Vancouver;
CAX;
FEATURES
source location/Qualifiers
1..1141
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 1.5%; Score 54; DB 6; Length 1141;
Best Local Similarity 10.0%; Pred. No. 0.45;
Matches 79; Conservative 311; Mismatches 389; Indels 8; Gaps 2;
2393 TTAGTGATGATCCCTGTGTGTAACAATGTAGCGCGTGTGAGCGCTGTGCTTTATG 2452
67 KYRRYNNKSRWKKWKKKKWBCANNYSBRYHARRKDKMTAYBMTNNKGTGMHR 126
2453 CAGAACTGTAGAACCCTCTGAA-AACTTGAACTAATGATTATGCAATTCGAGCTGG 2511
127 YMRWAMBIDYDHHVYTAANNAMWTTMCMDKDRTTRWKKNNATGMDDDTKYHMMNN 186
2512 GTCTCAGAGATGACGCTTGATGAAACTACTGTGTAAGTCCGTTGCTGATATGTTAAAG 2571
187 NGBCTVTWVVRKYTRDMSBKRMNYGMBWKNWSYDVYVWVMDMCKRYRRAVVRTRG 246
2572 GAGGCAAGTGAAGATCTGCTGTGTGAGCAATGATGATTTGATTTCACTGTTCAGC 2631
247 RRRNYMAMBTAHRRYNNGMTBMAAYRRWTTNNNNNNNAKMKCRKATGMWRABVNSTC 306
2632 CAGAAGTATTTCTTAAAGACGCTCATCTTTTCAACGCAAGATATGCTTTCTATG 2691
307 TTWKSRTTKVTSQVANNCRAGDANKDHKMKMSAAGVYNNNNNNNNWTKKARHBAW 366
2692 GAATCTGATGCTGACCATAGATGATGATTAATGATGCAATTTCAATATCTGCAATG 2751
367 DMVHSAWKMHANAHAHYSRKMTBYKRKTIVNNNNNGTTMKRMAMWYKMDMBGTYN 426
2752 CTCAAATATGCTGTGTTTGTGAGCTAAGAACTAGTCCCACTTAATACATGTCGCAA 2811
427 NNNNGRTTYGWTNKKKMTTYKKRANNCCKRAMDHKTCHNNTTWKMKTYMNNCYWKS 486
2812 AGTTGTACCAAGATTAAACAAGTTGCTGAGTAATTTCACTAATTAATGCT-----GCTT 2864
487 MTNGKSHRBAALVYTWMMWRVYAHANNNNNDYWKACITWYKVBCKMMNNVAAWTK 546
2865 GAATTTTGTGATCAACCTGTAGACAGAAATGTAATTCACTCACTCAATTTCTGTTAG 2924
547 SSMWYTSRYRWKTNNSWRMSDTRSMBRANNVYARABHYGKMNTRWMBWBSHTWBHBAAG 606
2925 AATAACGTAGATGATGAGATGCTTAGTGCTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGT 2984
607 AAHYWMBWYBAKCHCKAMYKAKKYAGAGSSNNNNNNNNNNNNNNNNNNATCARDYYAAS 666
2985 TTTTTCCTTTTCATTTAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3044
667 RRYAANAANKVYVYKBAANNAVYTHANNMWCNNNAATDTRTWKNNNNNNAGTWKNNNN 726
3045 ATGCTAGACGCTGACAGATATAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3104
727 NAKNSAAKNVYAAVYKAKKRWANVYKMGWADAAABTTDCKRNGAYKYKTYTNNN 786
3105 GGCCTGATCAGCGCATAGAAATGTTACCGAGGTGAGGGAATTAATCTCAATTCATCA 3164
787 NTYRGVYVNTAARDGANNAY 846

ORIGIN
Query Match 1.5%; Score 54; DB 6; Length 1141;
Best Local Similarity 10.0%; Pred. No. 0.45;
Matches 79; Conservative 311; Mismatches 389; Indels 8; Gaps 2;
2393 TTAGTGATGATCCCTGTGTGTAACAATGTAGCGCGTGTGAGCGCTGTGCTTTATG 2452
67 KYRRYNNKSRWKKWKKKKWBCANNYSBRYHARRKDKMTAYBMTNNKGTGMHR 126
2453 CAGAACTGTAGAACCCTCTGAA-AACTTGAACTAATGATTATGCAATTCGAGCTGG 2511
127 YMRWAMBIDYDHHVYTAANNAMWTTMCMDKDRTTRWKKNNATGMDDDTKYHMMNN 186
2512 GTCTCAGAGATGACGCTTGATGAAACTACTGTGTAAGTCCGTTGCTGATATGTTAAAG 2571
187 NGBCTVTWVVRKYTRDMSBKRMNYGMBWKNWSYDVYVWVMDMCKRYRRAVVRTRG 246
2572 GAGGCAAGTGAAGATCTGCTGTGTGAGCAATGATGATTTGATTTCACTGTTCAGC 2631
247 RRRNYMAMBTAHRRYNNGMTBMAAYRRWTTNNNNNNNAKMKCRKATGMWRABVNSTC 306
2632 CAGAAGTATTTCTTAAAGACGCTCATCTTTTCAACGCAAGATATGCTTTCTATG 2691
307 TTWKSRTTKVTSQVANNCRAGDANKDHKMKMSAAGVYNNNNNNNNWTKKARHBAW 366
2692 GAATCTGATGCTGACCATAGATGATGATTAATGATGCAATTTCAATATCTGCAATG 2751
367 DMVHSAWKMHANAHAHYSRKMTBYKRKTIVNNNNNGTTMKRMAMWYKMDMBGTYN 426
2752 CTCAAATATGCTGTGTTTGTGAGCTAAGAACTAGTCCCACTTAATACATGTCGCAA 2811
427 NNNNGRTTYGWTNKKKMTTYKKRANNCCKRAMDHKTCHNNTTWKMKTYMNNCYWKS 486
2812 AGTTGTACCAAGATTAAACAAGTTGCTGAGTAATTTCACTAATTAATGCT-----GCTT 2864
487 MTNGKSHRBAALVYTWMMWRVYAHANNNNNDYWKACITWYKVBCKMMNNVAAWTK 546
2865 GAATTTTGTGATCAACCTGTAGACAGAAATGTAATTCACTCACTCAATTTCTGTTAG 2924
547 SSMWYTSRYRWKTNNSWRMSDTRSMBRANNVYARABHYGKMNTRWMBWBSHTWBHBAAG 606
2925 AATAACGTAGATGATGAGATGCTTAGTGCTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGT 2984
607 AAHYWMBWYBAKCHCKAMYKAKKYAGAGSSNNNNNNNNNNNNNNNNNNNNNNNNNNNNATCARDYYAAS 666

OY 2985 TTTTCTTTTCATTAGGTCAGTCAAGCTGACGATTCAGAGCACTCCAGATG 3044
DB 667 RMYAMANAAMWYKBAANNAAYTTHANNMWGCNNNATDTRRTWKNNNNNAGTWKNNNN 726
OY 3045 ATGCTAGACTGCAGAGATATAGTATCCAGTGCAGAGATTAACTCTGCTTTG 3104
DB 727 NAKNAAKAAKAAVKAARKRWANRWAMRGWADAAABTTDKRNNGAYTKYTTTNN 786
OY 3105 GGCTGATCACCGCATAGAAATGTTACCGAGGTGAGGAGATTAATCTACAATCAATCA 3164
DB 787 NTYRGVITTAARDGMANNNNNNNNNNNNNGWSDWVTWAAAYGTNNNNNNNNNAY 846
OY 3165 ATTGTT 3171
DB 847 AMWTNKW 853

Search completed: December 11, 2005, 03:43:31
Job time : 18184.9 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 10, 2005, 07:52:23 ; Search time 204.429 Seconds
(without alignments)
4400.037 Million cell updates/sec

Title: US-10-600-070b-1

Perfect score: 2406
Sequence: 1 atggaagctcgcagctacacgt.....gctctgtcttcgcatcataa 2406

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3392430 seqs, 186927314 residues

Total number of hits satisfying chosen parameters: 6784860

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA_New:

1: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
2: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
5: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	44.8	1.9	191684	US-11-121-086-2	Sequence 2, Appli
C 2	43	1.8	1281	US-10-131-826A-509	Sequence 509, App
C 3	42.4	1.8	169495	US-11-121-086-61	Sequence 61, Appli
4	40.2	1.7	31028	US-10-829-826B-21	Sequence 21, Appli
5	40.2	1.7	31028	US-10-829-826B-22	Sequence 22, Appli
6	40.2	1.7	31028	US-10-829-826B-26	Sequence 26, Appli
7	40.2	1.7	31100	US-10-829-826B-24	Sequence 24, Appli
8	39.8	1.7	1450	US-11-140-417-3	Sequence 3, Appli
9	38.8	1.7	168516	US-11-121-086-3	Sequence 10, Appli
C 10	38.4	1.6	199321	US-11-121-086-10	Sequence 1, Appli
C 11	39.2	1.6	1317	US-11-140-417-1	Sequence 22, Appli
C 12	39.2	1.6	11462	US-11-140-417-22	Sequence 22, Appli
C 13	39.2	1.6	120096	US-11-121-086-24	Sequence 5, Appli
C 14	39	1.6	153376	US-11-121-086-5	Sequence 23, Appli
15	38.8	1.6	1772	US-11-167-856-23	Sequence 2713, A
16	38.8	1.6	7402	US-10-750-185-27313	Sequence 10, Appli
C 17	38.6	1.6	126552	US-11-121-086-1	Sequence 2713, A
C 18	38.2	1.6	7402	US-10-750-185-27313	Sequence 1058, Ap
C 19	38	1.6	17520	US-11-108-172-1058	Sequence 53, Appli
20	38	1.6	176503	US-11-121-086-53	Sequence 123, App
C 21	37.8	1.6	6840	US-11-090-739-123	Sequence 3097, Ap
C 22	37.6	1.6	600	US-10-750-185-3097	Sequence 25, Appli
C 23	37.6	1.6	173602	US-11-121-086-25	

24	37	1.5	31032	US-10-829-826B-23	Sequence 23, Appli
25	37	1.5	126552	US-11-121-086-1	Sequence 1, Appli
26	37	1.5	164810	US-11-121-086-4	Sequence 4, Appli
27	37	1.5	191684	US-11-121-086-2	Sequence 2, Appli
C 28	36.8	1.5	1044	US-10-508-263-109	Sequence 109, App
C 29	36.8	1.5	1213	US-11-112-908-420	Sequence 420, App
C 30	36.8	1.5	3880	US-10-485-517-9	Sequence 9, Appli
C 31	36.8	1.5	177623	US-11-112-908-41	Sequence 41, Appli
C 32	36.6	1.5	101786	US-11-117-187-199	Sequence 199, App
C 33	36.6	1.5	200628	US-11-121-086-62	Sequence 62, Appli
C 34	36	1.5	167116	US-11-121-086-44	Sequence 44, Appli
C 35	35.8	1.5	2492	US-10-821-234-629	Sequence 629, App
C 36	35.6	1.5	513	US-10-750-185-2594	Sequence 2594, Ap
37	35.4	1.5	1538	US-10-750-185-61710	Sequence 61710, A
C 38	35	1.5	702	US-10-793-626-735	Sequence 735, Appli
C 39	35	1.5	1557	US-10-996-217A-6	Sequence 6, Appli
C 40	35	1.5	3581	US-10-793-626-3817	Sequence 3817, Ap
C 41	34.8	1.4	131855	US-11-112-908-29	Sequence 29, Appli
C 42	34.8	1.4	143389	US-11-112-908-30	Sequence 30, Appli
C 43	34.8	1.4	150173	US-11-112-908-26	Sequence 26, Appli
C 44	34.8	1.4	166020	US-11-112-908-28	Sequence 28, Appli
C 45	34.8	1.4	171247	US-11-112-908-27	Sequence 27, Appli

ALIGNMENTS

```
RESULT 1
US-11-121-086-2/c
; Sequence 2, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; TITLE OR INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138, 6000-00000
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 2
; LENGTH: 191684
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-2

Query Match 1.9%; Score 44.8; DB 7; Length 191684;
Best Local Similarity 60.8%; Pred. No. 0.088;
Matches 73; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 85 AGCCACACACCTTACATCTATGCTCCGCGCAAGATGAGCCGCTCTCTCC 144
Db 50563 ACCCTCACCGGACAGACCTCCATCTCCATCTCCACCTCCAGATCTCTCC 50504
QY 145 GACTTCATTTACCTCGATTTCTCTCTCTCTCTCTGCGCACCGACCGCC 204
Db 50503 TCCATCACCTCCATCTCCATCTCCACCTCCAGACCTCCAGACCTCCAC 50444

RESULT 2
US-10-131-826A-509/c
; Sequence 509, Application US/10131826A
; Publication No. US20050245730A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
```

```

; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zhen
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C128
; CURRENT APPLICATION NUMBER: US/10/131,826A
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 509
; LENGTH: 1281
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-131-826A-509

Query Match      1.8%; Score 43; DB 6; Length 1281;
Best Local Similarity 47.6%; Pred. No. 0.0033;
Matches 127; Conservative 0; Mismatches 140; Indels 0; Gaps 0;

QY 421 TCTATCTCGGTCTAGAGAGAGATGACATGAGGCTCTTGTGATGATGAGAGAGCTTCA 480
DB 581 TTTGATCTCTTTCTTTGCTTCAAGAAATGCAATGTTGTTGATTTGTTACTGATGAGAGACACA 522
QY 481 GTCATCACTGATGTTCTTTGGATTAAGGTTCTTGGGGCTCTCTGTATGTCAGAGAGAGCT 540
DB 521 GATGTCATCTGACATGTTGTGGGTTACGGTCATGTTGATGTTGATATCTGAGATGTGTTTC 462
QY 541 GGTGAGACTGAGATAGTTCTTCTGGGTGTGAGGCTCTGCTTAAGAGAGGTTGCTTAAG 600
DB 461 TGTGAAACACTGTTGTTTGGGTGTAGACTTTAAGGTGATGATGATATTTGTTGAG 402
QY 601 TCGTTAAGCAAGATGTGTTTATGTTATGCGGCTTGGCTTCTGATGTTCTCGAGGGAT 660
DB 401 ACCATCCCTGTTGTTGTTGTTATTAATGCGGCTGTAAGTTTCATGTTGATGATGCGTTGTA 342
QY 661 GCTATGCAATGATGATCAACCTGATTTT 687
DB 341 TTACTGAGATCTGAGGCAACTGAACTT 315

RESULT 3
US-11-121-086-61/c
; Sequence 61, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
```

```

; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; PRIOR FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 61
; LENGTH: 169495
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (70072)..(70171)
; OTHER INFORMATION: a, c, g, t, unknown or other
; FEATURE:
; NAME/KEY: modified base
; LOCATION: (139457)..(157244)
; OTHER INFORMATION: a, c, g, t, unknown or other
US-11-121-086-61

Query Match      1.8%; Score 42.4; DB 7; Length 169495;
Best Local Similarity 48.7%; Pred. No. 0.44;
Matches 115; Conservative 0; Mismatches 121; Indels 0; Gaps 0;

QY 1657 GTCTTTAGTGTAGTCTCTGTTTGTGTAACATGTAGCCGTATGTAGCCCTGATGCTTT 1716
DB 67566 GTGGTGTCTGCTGATGCTGTTGATGATGATGATGATGATGATGATGATGATGATGATG 67507
QY 1717 ATTGCAAGCTGTGAAGACCTCTGAAACTTGAACATGATTTATGCAATTCAGACT 1776
DB 67506 ACTGTGTGTTGTTTACTCTGTTGTTGATGATGATGATGATGATGATGATGATGATGAT 67447
QY 1777 GGGGTCTCAAGAGTACGTTGATGATGAACACTGTTGAAATGTCCTGCTGATATGTTA 1836
DB 67446 GCTGTGATGATGCTGTTGTTGATGATGCTGCTGTTGTTGATGATGCTGCTGTTGTTG 67387
QY 1837 AAGAGCAAGTGTGAAGATCTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1892
DB 67386 ATTGCTCTTTTGTGTTGTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 67331

RESULT 4
US-10-829-826B-21
; Sequence 21, Application US/10829826B
; Publication No. US20050266397A1
; GENERAL INFORMATION:
; APPLICANT: Ecker, David J.
; APPLICANT: Hofstadler, Steven A.
; APPLICANT: Sampath, Rangarajan
; APPLICANT: Blyn, Lawrence B.
; APPLICANT: Hall, Thomas A.
; APPLICANT: Massire, Christian
; TITLE OF INVENTION: METHOD FOR IDENTIFICATION OF CORONAVIRUSES
; FILE REFERENCE: IBIS0075-100 (DIBIS-0058US)
; CURRENT APPLICATION NUMBER: US/10/829,826B
; CURRENT FILING DATE: 2004-04-22
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 31028
; TYPE: DNA
; ORGANISM: Coronavirus
US-10-829-826B-21

Query Match      1.7%; Score 40.2; DB 6; Length 31028;
Best Local Similarity 51.4%; Pred. No. 0.63;
Matches 93; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 1651 GAGACAGTGTATTAGTGTAGATCTGTTGTTGTAACAATGTAGCCGCTGATGTGAGCTGCT 1710
```

Db 3700 GATGCTATGTTCTTTTATGATGATGTTGTCACATGTCGACAGTGTGATGCTATG 3759
QY 1711 GTCTTTATTCAGAGAGCTTAAGACCCCTGTAACCTTGAACCTAATGATTAATGCAAT 1770
Db 3760 GTACTTATTTAGTGTGATGATGTCATTTACAGCCCACTTGTCTTTAAGATTAAGTTGTT 3819
QY 1771 CGAGCTGGGGTCTCAGAGAGTACGCTTGATGTAAGTGAATCTGTTGAATGTCCGTTGCTGAT 1830
Db 3820 TGTGCAATTTATTAACGAGCGTATGTTGTAATTAAGCAGCTTGTGTTGATGATTAATGAT 3879
QY 1831 A 1831
Db 3880 A 3880

RESULT 5
US-10-829-826B-22

; Sequence 22, Application US/10829826B
; Publication No. US20050266397A1
; GENERAL INFORMATION:
; APPLICANT: Ecker, David J.
; APPLICANT: Hofstadler, Steven A.
; APPLICANT: Sampath, Rangarajan
; APPLICANT: Blyn, Lawrence B.
; APPLICANT: Hall, Thomas A.
; APPLICANT: Masire, Christian
; TITLE OF INVENTION: METHOD FOR IDENTIFICATION OF CORONAVIRUSES
; FILE REFERENCE: IBIS0075-100 (DIBIS-0058US)
; CURRENT FILING DATE: 2004-04-22
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 22
; LENGTH: 31028
; TYPE: DNA
; ORGANISM: Coronavirus
US-10-829-826B-22

Query Match 1.7%; Score 40.2; DB 6; Length 31028;

Best Local Similarity 51.4%; Pred. No. 0.63; Matches 93; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 1651 GAGACAGTGTAGTATGATCTGTTGTAACAATGTAAGCCGTGATGAGCCCTGCT 1710
Db 3700 GATGCTATGTTCTTTTATGATGATGTTGTCACATGTCGACAGTGTGATGCTATG 3759
QY 1711 GTCTTTATTCAGAGAGCTTAAGACCCCTGTAACCTTGAACCTAATGATTAATGCAAT 1770
Db 3760 GTACTTATTTAGTGTGATGATGTCATTTACAGCCCACTTGTCTTTAAGATTAAGTTGTT 3819
QY 1771 CGAGCTGGGGTCTCAGAGAGTACGCTTGATGTAAGTGAATCTGTTGAATGTCCGTTGCTGAT 1830
Db 3820 TGTGCAATTTATTAACGAGCGTATGTTGTAATTAAGCAGCTTGTGTTGATGATTAATGAT 3879
QY 1831 A 1831
Db 3880 A 3880

RESULT 6
US-10-829-826B-26

; Sequence 26, Application US/10829826B
; Publication No. US20050266397A1
; GENERAL INFORMATION:
; APPLICANT: Ecker, David J.
; APPLICANT: Hofstadler, Steven A.
; APPLICANT: Sampath, Rangarajan
; APPLICANT: Blyn, Lawrence B.
; APPLICANT: Hall, Thomas A.
; APPLICANT: Masire, Christian
; TITLE OF INVENTION: METHOD FOR IDENTIFICATION OF CORONAVIRUSES
; FILE REFERENCE: IBIS0075-100 (DIBIS-0058US)

; CURRENT APPLICATION NUMBER: US/10/829, 826B
; CURRENT FILING DATE: 2004-04-22
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 26
; LENGTH: 31028
; TYPE: DNA
; ORGANISM: Coronavirus
US-10-829-826B-26

Query Match 1.7%; Score 40.2; DB 6; Length 31028;

Best Local Similarity 51.4%; Pred. No. 0.63; Matches 93; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 1651 GAGACAGTGTAGTATGATCTGTTGTAACAATGTAAGCCGTGATGAGCCCTGCT 1710
Db 3700 GATGCTATGTTCTTTTATGATGATGTTGTCACATGTCGACAGTGTGATGCTATG 3759
QY 1711 GTCTTTATTCAGAGAGCTTAAGACCCCTGTAACCTTGAACCTAATGATTAATGCAAT 1770
Db 3760 GTACTTATTTAGTGTGATGATGTCATTTACAGCCCACTTGTCTTTAAGATTAAGTTGTT 3819
QY 1771 CGAGCTGGGGTCTCAGAGAGTACGCTTGATGTAAGTGAATCTGTTGAATGTCCGTTGCTGAT 1830
Db 3820 TGTGCAATTTATTAACGAGCGTATGTTGTAATTAAGCAGCTTGTGTTGATGATTAATGAT 3879
QY 1831 A 1831
Db 3880 A 3880

RESULT 7
US-10-829-826B-24

; Sequence 24, Application US/10829826B
; Publication No. US20050266397A1
; GENERAL INFORMATION:
; APPLICANT: Ecker, David J.
; APPLICANT: Hofstadler, Steven A.
; APPLICANT: Sampath, Rangarajan
; APPLICANT: Blyn, Lawrence B.
; APPLICANT: Hall, Thomas A.
; APPLICANT: Masire, Christian
; TITLE OF INVENTION: METHOD FOR IDENTIFICATION OF CORONAVIRUSES
; FILE REFERENCE: IBIS0075-100 (DIBIS-0058US)
; CURRENT APPLICATION NUMBER: US/10/829, 826B
; CURRENT FILING DATE: 2004-04-22
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 24
; LENGTH: 31100
; TYPE: DNA
; ORGANISM: Coronavirus
US-10-829-826B-24

Query Match 1.7%; Score 40.2; DB 6; Length 31100;

Best Local Similarity 51.4%; Pred. No. 0.63; Matches 93; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 1651 GAGACAGTGTAGTATGATCTGTTGTAACAATGTAAGCCGTGATGAGCCCTGCT 1710
Db 3700 GATGCTATGTTCTTTTATGATGATGTTGTCACATGTCGACAGTGTGATGCTATG 3759
QY 1711 GTCTTTATTCAGAGAGCTTAAGACCCCTGTAACCTTGAACCTAATGATTAATGCAAT 1770
Db 3760 GTACTTATTTAGTGTGATGATGTCATTTACAGCCCACTTGTCTTTAAGATTAAGTTGTT 3819
QY 1771 CGAGCTGGGGTCTCAGAGAGTACGCTTGATGTAAGTGAATCTGTTGAATGTCCGTTGCTGAT 1830
Db 3820 TGTGCAATTTATTAACGAGCGTATGTTGTAATTAAGCAGCTTGTGTTGATGATTAATGAT 3879
QY 1831 A 1831
Db 3880 A 3880

APPLICANT: Ke, Jinsan
APPLICANT: Johnson, Jerry L
APPLICANT: Allred, Carolyn C
APPLICANT: Farland, Beth
APPLICANT: Lutziger, Isabelle
APPLICANT: Wen, Teu-Jung
TITLE OF INVENTION: Materials and Methods for the Alteration of Enzyme and
FILE OF INVENTION: Acetyl CoA Levels in Plants
FILE REFERENCE: P2194USDIV-2
CURRENT APPLICATION NUMBER: US/11/167,856
CURRENT FILING DATE: 2005-06-27
PRIOR APPLICATION NUMBER: US 10/293,865
PRIOR FILING DATE: 2002-11-13
PRIOR APPLICATION NUMBER: US 09/344,882
PRIOR FILING DATE: 1999-06-25
PRIOR APPLICATION NUMBER: US 60/090,717
PRIOR FILING DATE: 1998-06-26
NUMBER OF SEQ ID NOS: 38
SOFTWARE: PatentIn Ver. 3.1
SEQ ID NO 23
LENGTH: 1772
TYPE: DNA
ORGANISM: Arabidopsis Thaliana
FEATURE:
NAME/KEY: exon
LOCATION: (1)..(834)
FEATURE:
NAME/KEY: exon
LOCATION: (836)..(930)
FEATURE:
NAME/KEY: exon
LOCATION: (932)..(939)
FEATURE:
NAME/KEY: exon
LOCATION: (941)..(1772)
US-11-167-856-23

Query Match 1.6%; Score 38.8; DB 7; Length 1772;
Best Local Similarity 46.6%; Pred. No. 0.23;
Matches 124; Conservative 0; Mismatches 142; Indels 0; Gaps 0;

QY	832	CCGCTTGATGATGATTACCGCGGAAAGACTAATGTTTAAAGCGGTGTGGGAATATT	891
DB	705	CCTCTATCTGCTCTTCTGTGGGAACTACTCATGAGGCTGACTTCTGATGAGATT	764
QY	892	TTGTGCTGTGTGAGAGAGTGAGCATCAGCTTTGTTGGGGTTTGACCGTGAGAAG	951
DB	765	GTGAATATAGTTTCTGGAATTTGGGCTACTGCTGTGACAGCTATAGTACATGAGC	824
QY	952	TTTATGAATGAGCGCTTTTACGAATGACAGCTGCTGAGCAGGTTGATCTTTTGTAGCT	1011
DB	825	GTTGATTAAGGTGCTTTCACCGGCTACTGATGTTGGGAAGATTATCTTGAGTTAGCT	884
QY	1012	ACCCAGCAATATTCAGCAAGTCAATTGAGTTTGAAGTTTGAAGTTGCACTTGTCTTGTG	1071
DB	885	TCAAAAAGCAACCTTAAGCAGTCACTTGAAGCTGAGGAAGTCACCAATTCATTGTGA	944
QY	1072	GCTCAAGCTTTTATTTGTTAAGGCC	1097
DB	945	TGTGAAGATGCTGATGTGATGAGGC	970

Search completed: December 10, 2005, 18:24:03
Job time : 208.429 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using SW model

Run on: December 10, 2005, 06:08:12 ; Search time 1835.5 Seconds
(without alignments)
10839.614 Million cell1 updates/sec

Title: US-10-600-070b-1

Perfect score: 2406
1 atcgaagctctgagtcacgt.....gctctgtcttcgcatcataa 2406

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_NA_Main:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2406	100.0	2406	US-10-600-070-1	Sequence 1, Appli
2	2402.8	99.9	2406	US-10-600-070-9	Sequence 9, Appli
3	2399.6	99.7	2406	US-10-600-070-128	Sequence 128, App
4	2399.6	99.7	2637	US-10-600-070-130	Sequence 130, App
5	2399.6	99.7	2679	US-10-739-930-227	Sequence 227, App
6	1724	71.7	3667	US-10-600-070-3	Sequence 3, Appli
7	1720.8	71.5	3667	US-10-600-070-10	Sequence 10, Appli
8	482	20.0	561	US-10-600-070-132	Sequence 132, App
9	481	20.0	2283	US-10-600-070-126	Sequence 126, App
10	472.8	19.7	1146	US-10-424-599-129007	Sequence 129007,
11	361.4	15.0	1411	US-10-425-115-81853	Sequence 81853, A
12	328.8	13.7	631	US-10-600-070-184	Sequence 184, App
13	278.8	11.6	660	US-10-600-070-135	Sequence 135, App
14	231.6	9.6	1039	US-10-424-599-35059	Sequence 35059, A
15	224	9.3	537	US-10-600-070-143	Sequence 143, App
16	223.4	9.2	552	US-10-021-323-3636	Sequence 3636, App
17	207	8.6	545	US-10-600-070-175	Sequence 175, App
18	197.2	8.2	491	US-10-600-070-114	Sequence 174, App
19	195.8	8.1	467	US-10-600-070-187	Sequence 185, App
20	193.8	8.1	460	US-10-600-070-187	Sequence 187, App
21	190	7.9	446	US-10-600-070-151	Sequence 151, App
22	188.8	7.8	1703	US-10-437-963-69933	Sequence 69933, A
23	186.6	7.8	652	US-10-600-070-186	Sequence 186, App

24	178.6	7.4	607	7	US-10-021-323-3562	Sequence 3562, App
25	175	7.3	439	3	US-09-732-627A-2154	Sequence 2154, App
26	170.4	7.1	608	7	US-10-600-070-137	Sequence 137, App
27	170.2	7.1	563	7	US-10-600-070-178	Sequence 178, App
28	169.8	7.1	622	7	US-10-600-070-146	Sequence 146, App
29	169	7.0	1536	8	US-10-425-115-57452	Sequence 57452, A
30	167.6	7.0	1032	7	US-10-767-701-9474	Sequence 9474, App
31	165.4	6.9	527	7	US-10-600-070-148	Sequence 148, App
32	162	6.7	647	7	US-10-767-701-4069	Sequence 4069, App
33	158.6	6.6	871	7	US-10-600-070-153	Sequence 153, App
34	155.4	6.5	307	7	US-10-600-070-138	Sequence 138, App
35	150.4	6.3	418	7	US-10-600-070-144	Sequence 144, App
36	146.2	6.1	2130	7	US-10-437-963-69932	Sequence 69932, A
37	144.6	6.0	653	7	US-10-600-070-148	Sequence 148, App
38	144.2	6.0	479	7	US-10-600-070-150	Sequence 150, App
39	130.8	5.4	420	7	US-10-600-070-176	Sequence 176, App
40	128.6	5.3	480	7	US-10-600-070-145	Sequence 145, App
41	127.6	5.3	535	7	US-10-600-070-149	Sequence 149, App
42	124.2	5.2	360	7	US-10-600-070-179	Sequence 179, App
43	118.8	4.9	336	7	US-10-600-070-142	Sequence 142, App
44	118.6	4.9	541	7	US-10-600-070-154	Sequence 154, App
45	117.8	4.9	309	7	US-10-600-070-141	Sequence 141, App

ALIGNMENTS

RESULT 1
US-10-600-070-1
; Sequence 1, Application US/10600070
; Publication No. US20040139500A1
; GENERAL INFORMATION:
; APPLICANT: Osteoryoung, Katherine W.
; APPLICANT: Vitcha, Stanislaw
; APPLICANT: Kokscharova, Olga A.
; TITLE OF INVENTION: Placid Division and Related Genes and Proteins, and Methods of
; FILE REFERENCE: MSU-08153
; CURRENT APPLICATION NUMBER: US/10/600,070
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 2406
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-600-070-1

Query Match	100.0%;	Score 2406;	DB 7;	Length 2406;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2406;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ATGGAAGCTCTGAGTCACTGTCGATTTGCTTCCCATTCGAATTAATGCGATTACCA	60	
DB	1	ATGGAAGCTCTGAGTCACTGTCGATTTGCTTCCCATTCGAATTAATGCGATTACCA	60	
QY	61	CCGGCGACGAACAAAGCTCCGACGTAAGCCAAACCTCTAACAATATCTGTCGGCAGC	120	
DB	61	CCGGCGACGAACAAAGCTCCGACGTAAGCCAAACCTCTAACAATATCTGTCGGCAGC	120	
QY	121	AAATGGGCGGACCGCTCTCTCTCCGACTTCAATTTCACTCCGATTCCTCTCTCTCC	180	
DB	121	AAATGGGCGGACCGCTCTCTCTCCGACTTCAATTTCACTCCGATTCCTCTCTCTCC	180	
QY	181	TTTCCGACCGGACCAACACCGCCACTCTGCTCTCTGCGACCATTAATGATGATGCTCC	240	
DB	181	TTTCCGACCGGACCAACACCGCCACTCTGCTCTCTGCGACCATTAATGATGATGCTCC	240	
QY	241	GAAAGCCACGCGCCATTCATTTGATTTGACAGGATTAAGAGCTCAACAACATTTTC	300	
DB	241	GAAAGCCACGCGCCATTCATTTGATTTGACAGGATTAAGAGCTCAACAACATTTTC	300	

QY 301 TTAAACGATGGAATCAGAAAGACATTCGAAAGCTAGGCTTTCGAAACCGCCGCAATTCGGT 360
Db 301 TTAAACCGATGGAATCAGAAAGACATTCGAAAGCTAGGCTTTCGAAACCGCCGCAATTCGGT 360
QY 361 TTCGCCGACGACGCTTTAATCAGCCGGAGACAGATTTCTTCAAGCTGCTTGGAAACCTCTG 420
Db 361 TTCGCCGACGACGCTTTAATCAGCCGGAGACAGATTTCTTCAAGCTGCTTGGAAACCTCTG 420
QY 421 TCTAATCCTCGGTCTAGAAAGAGATACAAATGAAAGCTTCTTGATGATGAAGAAGCTACA 480
Db 421 TCTAATCCTCGGTCTAGAAAGAGATACAAATGAAAGCTTCTTGATGATGAAGAAGCTACA 480
QY 481 GTCACTACTGATGTTCTTGAGATTAAGGTTCTCTGAGGCTCTCTGTGTATTCGAAAGAGT 540
Db 481 GTCACTACTGATGTTCTTGAGATTAAGGTTCTCTGAGGCTCTCTGTGTATTCGAAAGAGT 540
QY 541 GGTGAGACTGAGATGTTCTTGAGGATTAAGGTTCTCTGAGGCTCTCTGTGTATTCGAAAGAGT 600
Db 541 GGTGAGACTGAGATGTTCTTGAGGATTAAGGTTCTCTGAGGCTCTCTGTGTATTCGAAAGAGT 600
QY 601 TCGTTTAAGCAAGATGTTGATTAAGGTTCTTGAGGCTCTCTGATGCTCGAGGAT 660
Db 601 TCGTTTAAGCAAGATGTTGATTAAGGTTCTTGAGGCTCTCTGATGCTCGAGGAT 660
QY 661 GCTATGCAATGGATCCACTGATTTAATCTGTTAAGTGTGAGGTTGTTGAGGAACTTTG 720
Db 661 GCTATGCAATGGATCCACTGATTTAATCTGTTAAGTGTGAGGTTGTTGAGGAACTTTG 720
QY 721 AAGCTTTTACAGAGAGAGAGCAAGTACCTTGACCGGATTTACGTCGCAAAATGAT 780
Db 721 AAGCTTTTACAGAGAGAGAGCAAGTACCTTGACCGGATTTACGTCGCAAAATGAT 780
QY 781 GAGAATTGGAAGAGATCACTCCGCTTATGCTTGGAGCTACTTGGCTTACCGCTTGGT 840
Db 781 GAGAATTGGAAGAGATCACTCCGCTTATGCTTGGAGCTACTTGGCTTACCGCTTGGT 840
QY 841 GATGATTAACGCTGCGAAAAGACTAATGCTTTAAGCGGTGTGCGGAATATTTTGTGCT 900
Db 841 GATGATTAACGCTGCGAAAAGACTAATGCTTTAAGCGGTGTGCGGAATATTTTGTGCT 900
QY 901 GTTGGAGAGGTTGAGCATCAAGCTCTGTGAGGCTTGAACCCGAGAAAGTTATGAT 960
Db 901 GTTGGAGAGGTTGAGCATCAAGCTCTGTGAGGCTTGAACCCGAGAAAGTTATGAT 960
QY 961 GAGGCGTTTTACGATGACAGCTGCTGAGAGGTTGATCTTTTGTGATACCCCAAGC 1020
Db 961 GAGGCGTTTTACGATGACAGCTGCTGAGAGGTTGATCTTTTGTGATACCCCAAGC 1020
QY 1021 AATATTCAGACAGAGTCAATTGAAAGTTTACGAAGTTGCACTTGTCTTGTGCTCAAGCT 1080
Db 1021 AATATTCAGACAGAGTCAATTGAAAGTTTACGAAGTTTGCACTTGTCTTGTGCTCAAGCT 1080
QY 1081 TTTATTTGTTAAGAAAGCAACCTTTTACAGAGTGTGATTAAGCAATTCAGCAACTTCAG 1140
Db 1081 TTTATTTGTTAAGAAAGCAACCTTTTACAGAGTGTGATTAAGCAATTCAGCAACTTCAG 1140
QY 1141 CAGGCTAAGGTATGCTATGAGAGATTCCTGCGATGTTGTATGATACAGCAATTAATTGG 1200
Db 1141 CAGGCTAAGGTATGCTATGAGAGATTCCTGCGATGTTGTATGATACAGCAATTAATTGG 1200
QY 1201 GAGATGACCTTCGCTTGAAGAAAGGGAATCTGTGCACTGCTTATTAAGCAAGTTGATGAA 1260
Db 1201 GAGATGACCTTCGCTTGAAGAAAGGGAATCTGTGCACTGCTTATTAAGCAAGTTGATGAA 1260
QY 1261 TGCCGATATGTTGGGCTTGAAGAGATTCGAATTAAGCAATTAAGCAATTAAGTGTG 1320
Db 1261 TGCCGATATGTTGGGCTTGAAGAGATTCGAATTAAGCAATTAAGCAATTAAGTGTG 1320
QY 1321 GAGTTTGTGTTTGAAGAAATTCGAATTCGATGACATGATGATCTCCCTGACATATGCAAA 1380
Db 1321 GAGTTTGTGTTTGAAGAAATTCGAATTCGATGACATGATGATCTCCCTGACATATGCAAA 1380
QY 1381 TTGTTGAAACCTGTTGGCAGGGGTGTCTTTCTTAGGTTCAAGACACCAAAAGATTA 1440

Db 1381 TTGTTGAAACCTGTTGGCAGGGGTGTCTTTCTTAGGTTCAAGACACCAAAAGATTA 1440
QY 1441 AAATTTAAACCTGGGGGACTATATGATGATCTATGATGTTTGAAGTTTCTTGGAAAGTGT 1500
Db 1441 AAATTTAAACCTGGGGGACTATATGATGATCTATGATGTTTGAAGTTTCTTGGAAAGTGT 1500
QY 1501 GAGGTAGTTCAAGGTTCTCTTATGCTGCTGCAACTATGAGCAAGATTTGAGCCGAG 1560
Db 1501 GAGGTAGTTCAAGGTTCTCTTATGCTGCTGCAACTATGAGCAAGATTTGAGCCGAG 1560
QY 1561 CATGTGAAGCTAGTGTCTATGACAGGCACTGCAAGAAATTTTCTTCCGCTATACGAT 1620
Db 1561 CATGTGAAGCTAGTGTCTATGACAGGCACTGCAAGAAATTTTCTTCCGCTATACGAT 1620
QY 1621 AGAAACTCGCTGAAACCAAGAGATGTGCAAGAGACAGTGTGATGATGATGATGATGAT 1680
Db 1621 AGAAACTCGCTGAAACCAAGAGATGTGCAAGAGACAGTGTGATGATGATGATGATGAT 1680
QY 1681 AACATGTAAGCCGCTGATGATGAGCTGCTGCTTATATTCAGAAAGCTGTAGACCTCT 1740
Db 1681 AACATGTAAGCCGCTGATGATGAGCTGCTGCTTATATTCAGAAAGCTGTAGACCTCT 1740
QY 1741 GAAAACTTTGAAACTAATGATTAATGCAATTCGAGCTGGGCTCTCAGAGATGAGCTTGT 1800
Db 1741 GAAAACTTTGAAACTAATGATTAATGCAATTCGAGCTGGGCTCTCAGAGATGAGCTTGT 1800
QY 1801 GAAACTACTGTTGAAAGTCCGCTGCTGATATGTTAAAGAGGCAAGCTGGAAGATCTTA 1860
Db 1801 GAAACTACTGTTGAAAGTCCGCTGCTGATATGTTAAAGAGGCAAGCTGGAAGATCTTA 1860
QY 1861 GCTGCTGCTGCTGCAATGGAATGCACTGATTTCACTGCTTACCGCAGAAATTTTCTTAAAGC 1920
Db 1861 GCTGCTGCTGCTGCAATGGAATGCACTGATTTCACTGCTTACCGCAGAAATTTTCTTAAAGC 1920
QY 1921 AGCTCATCTTTTCAACCCAGATATGATGTTTCTTATGAAATCTGATGCTGCTACATA 1980
Db 1921 AGCTCATCTTTTCAACCCAGATATGATGTTTCTTATGAAATCTGATGCTGCTACATA 1980
QY 1981 GGGTCAGTCAAGAGCTGACATTCGAAAGCACTTCCGAAATGATGCTGAGACAGCT 2040
Db 1981 GGGTCAGTCAAGAGCTGACATTCGAAAGCACTTCCGAAATGATGCTGAGACAGCT 2040
QY 2041 AATATAGTATCCAAGTGGCAGAAAGATTAAGTCTCTGCTTTTGGGCTGATCAACCGATA 2100
Db 2041 AATATAGTATCCAAGTGGCAGAAAGATTAAGTCTCTGCTTTTGGGCTGATCAACCGATA 2100
QY 2101 GAAATGTTACCAAGGTTTGGATGGCGCAATGCTGAAGATTTGGAATCTGACAGAGCT 2160
Db 2101 GAAATGTTACCAAGGTTTGGATGGCGCAATGCTGAAGATTTGGAATCTGACAGAGCT 2160
QY 2161 GAAACTGCGAGCTGGGTTGTTATGATTAATACCTGTTGAAACTATCTGTTGACAGT 2220
Db 2161 GAAACTGCGAGCTGGGTTGTTATGATTAATACCTGTTGAAACTATCTGTTGACAGT 2220
QY 2221 GTGACAGTCTCAGCAGATGGAACCCGCTCTGCTGAGAACCAACTGAGAGAGTCTGCT 2280
Db 2221 GTGACAGTCTCAGCAGATGGAACCCGCTCTGCTGAGAACCAACTGAGAGAGTCTGCT 2280
QY 2281 TGTCTATCTGATTTGGTTTCAATCCAGAAACAAATGCTATCTGATGCTGACAACTACAA 2340
Db 2281 TGTCTATCTGATTTGGTTTCAATCCAGAAACAAATGCTATCTGATGCTGACAACTACAA 2340
QY 2341 AGATACGAAGTTTCTGCTCAGAGTCAAGGTTGAAATCACTGAAAGCTCTGTTCTTGCA 2400
Db 2341 AGATACGAAGTTTCTGCTCAGAGTCAAGGTTGAAATCACTGAAAGCTCTGTTCTTGCA 2400
QY 2401 TCATTA 2406
Db 2401 TCATTA 2406

RESULT 2

US-10-600-070-9
; Sequence 9, Application US/10600070
; Publication No. US20040139500A1
; GENERAL INFORMATION:
; APPLICANT: Oetereyoung, Katherine W.
; APPLICANT: Vittha, Stanislaw
; APPLICANT: Kokeharova, Olga A.
; APPLICANT: Gao, Hongbo
; TITLE OF INVENTION: Plasmid Division and Related Genes and Proteins, and Methods of
; TITLE OF INVENTION: Use
; FILE REFERENCE: MSU-08153
; CURRENT FILING DATE: 2003-06-20
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 9
; LENGTH: 2406
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-600-070-9

Query Match 99.9%; Score 2402.8; DB 7; Length 2406;
Best Local Similarity 99.9%; Pred No. 0;

Matches 2404; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGAAGCTCTGAGTCAAGTCGCGCATTTGGTCTCTGCCAATTCGAATTAATGCGGATTACCA 60
DB 1 ATGGAAGCTCTGAGTCAAGTCGCGCATTTGGTCTCTGCCAATTCGAATTAATGCGGATTACCA 60
QY 61 CCGGGAAGCAAAAGCTCCGAGTGGAGCAACACCTCTACAACTATCTGCTCCGCGAC 120
DB 61 CCGGGAAGCAAAAGCTCCGAGTGGAGCAACACCTCTACAACTATCTGCTCCGCGAC 120
QY 121 AAATGGGCGAGCCGCTTCTCTCCGACTTCAATTTCACTCCGATTCCTCCCTCTCTCC 180
DB 121 AAATGGGCGAGCCGCTTCTCTCCGACTTCAATTTCACTCCGATTCCTCCCTCTCTCC 180
QY 181 TTCCGCAACCGCCACCAACCAACCGCACTCTCGTCTCTGCCAATTAATGATGCTCC 240
DB 181 TTCCGCAACCGCCACCAACCAACCGCACTCTCGTCTCTGCCAATTAATGATGCTCC 240
QY 241 GAAGCCGCAAGTCCCAATCCCAATTTATTTCTACCGAGTATTAAGAGCTCAAAACATTTG 300
DB 241 GAAGCCGCAAGTCCCAATCCCAATTTATTTCTACCGAGTATTAAGAGCTCAAAACATTTG 300
QY 301 TTAACCGATGGAATCAGAAAGAGCATTCGAAGCTAGGTTTGAACCGCGCAATTCGGT 360
DB 301 TTAACCGATGGAATCAGAAAGAGCATTCGAAGCTAGGTTTGAACCGCGCAATTCGGT 360
QY 361 TTACGCGAGCAGCTTTTATATCAGCCGAGACAGATTTCTTCAAGCTGCTTGCGAAACTCTG 420
DB 361 TTACGCGAGCAGCTTTTATATCAGCCGAGACAGATTTCTTCAAGCTGCTTGCGAAACTCTG 420
QY 421 TCTAATCTCGGTCTAGAAAGAGATCAATGAAGGCTCTTCTGATGATGAAGAAGCTACA 480
DB 421 TCTAATCTCGGTCTAGAAAGAGATCAATGAAGGCTCTTCTGATGATGAAGAAGCTACA 480
QY 481 GTCATCATGATGTTCTTTGGGATAAGGTTCTGGGGCTCTCTGTATGTAAGTAAGAAGT 540
DB 481 GTCATCATGATGTTCTTTGGGATAAGGTTCTGGGGCTCTCTGTATGTAAGTAAGAAGT 540
QY 541 GGTGAGACTGAGATGTTCTTCGGGTTGGTAGGCTCTGCTTAAAGAGAGGTTGCTTAAG 600
DB 541 GGTGAGACTGAGATGTTCTTCGGGTTGGTAGGCTCTGCTTAAAGAGAGGTTGCTTAAG 600
QY 601 TCGTTAAAGCAAGATGTTTATGATATGAGGCTTGGCTTCTGATGATCTCGAGGGAT 660
DB 601 TCGTTAAAGCAAGATGTTTATGATATGAGGCTTGGCTTCTGATGATCTCGAGGGAT 660
QY 661 GCTATGCAATGATGCACTGATTTTATTAAGTATGATGATGATGATGATGATGATGAT 720
DB 661 GCTATGCAATGATGCACTGATTTTATTAAGTATGATGATGATGATGATGATGATGAT 720

QY 721 AAGCTTTTACAGAGAGAGCAAGTATGCTTGACCCGATTTTACGTCACAAATTCAT 780
DB 721 AAGCTTTTACAGAGAGAGCAAGTATGCTTGACCCGATTTTACGTCACAAATTCAT 780
QY 781 GAGACTTTGAAGAGATCATCCGCGTATGCTTGAAGCTACTGAGCTTCCGCTGCTGCT 840
DB 781 GAGACTTTGAAGAGATCATCCGCGTATGCTTGAAGCTACTGAGCTTCCGCTGCTGCT 840
QY 841 GATGATTAACGTCGCAAAAAGCTAAATGTTTAAAGCGGTGTCGGAATATTTTGTGCT 900
DB 841 GATGATTAACGTCGCAAAAAGCTAAATGTTTAAAGCGGTGTCGGAATATTTTGTGCT 900
QY 901 GTTGAGAGAGGTGAGCATTCAGCTCTTGTGGGGTTGACCCGTGAGAAATTAATGAT 960
DB 901 GTTGAGAGAGGTGAGCATTCAGCTCTTGTGGGGTTGACCCGTGAGAAATTAATGAT 960
QY 961 GAGCGCTTTTACAAATGACAGCTGTCGAGAGAGTGAATCTTTTGTATGATACCCCAAG 1020
DB 961 GAGCGCTTTTATGATATACAGCTGTCGAGAGAGTGAATCTTTTGTATGATACCCCAAG 1020
QY 1021 AATATTCAGAGAGTCAATTTGAAGTTTACGAAGTTCGACTTGTGCTGCTCAAGCT 1080
DB 1021 AATATTCAGAGAGTCAATTTGAAGTTTACGAAGTTCGACTTGTGCTGCTCAAGCT 1080
QY 1081 TTTATGCTTAAGAAACCAACCTTTTACAGATGCTGATTAAGCAATTCAGCAACTTCA 1140
DB 1081 TTTATGCTTAAGAAACCAACCTTTTACAGATGCTGATTAAGCAATTCAGCAACTTCA 1140
QY 1141 CAGGCTAAGTATAGGCTATGAGATTCCTGCGATGTTGATGATGATGATGATGATGAT 1200
DB 1141 CAGGCTAAGTATAGGCTATGAGATTCCTGCGATGTTGATGATGATGATGATGATGAT 1200
QY 1201 GAGATAGACTTCGCTCTAGAAAGGGGACTCTGTGCACTGCTTATAGGCAAAATTCATG 1260
DB 1201 GAGATAGACTTCGCTCTAGAAAGGGGACTCTGTGCACTGCTTATAGGCAAAATTCATG 1260
QY 1261 TGCCGATATGCTTGGGCTTGAACAAGTGAAGATTCACATATTAAGCAATTCAGCTAT 1320
DB 1261 TGCCGATATGCTTGGGCTTGAACAAGTGAAGATTCACATATTAAGCAATTCAGCTAT 1320
QY 1321 GAGTTGTTTGAAGAAATTCAAATCGTATGATGATGATGATGATGATGATGATGATG 1380
DB 1321 GAGTTGTTTGAAGAAATTCAAATCGTATGATGATGATGATGATGATGATGATGATG 1380
QY 1381 TTGTTGGAACCTGCTGTTGAGGAGGCTTCTTCTGATGCTCAGAGACCAACCAAGATTA 1440
DB 1381 TTGTTGGAACCTGCTGTTGAGGAGGCTTCTTCTGATGCTCAGAGACCAACCAAGATTA 1440
QY 1441 AAATTTAACTCGGAGACTATGATGATGATGATGATGATGATGATGATGATGATGATG 1500
DB 1441 AAATTTAACTCGGAGACTATGATGATGATGATGATGATGATGATGATGATGATGATG 1500
QY 1501 GAGGTATGTCAGGCTTCTCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560
DB 1501 GAGGTATGTCAGGCTTCTCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560
QY 1561 CATGTGAAGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTAT 1620
DB 1561 CATGTGAAGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTAT 1620
QY 1621 AGAAATCTCGCTGAACCCAGAGATGTCAGAGACAGTGTATGATGATGATGATGATG 1680
DB 1621 AGAAATCTCGCTGAACCCAGAGATGTCAGAGACAGTGTATGATGATGATGATGATG 1680
QY 1681 AACATGATGAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
DB 1681 AACATGATGAGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
QY 1741 GAAATCTTTGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800
DB 1741 GAAATCTTTGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800
QY 1801 GAAATCTGTTGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860

1801 GAAATCTACTGTTGAAATGCTCCGTTGCTGATATGTTAAAGGAGGCAAGTGTGAAGTCTTA 1860
1861 GCTGCTGGTGGCAATTTGACATGATTTCACTGTTCAAGCAAGATATTTCTTTAAAGC 1920
1861 GCTGCTGGTGGCAATTTGACATGATTTCACTGTTCAAGCAAGATATTTCTTTAAAGC 1920
1921 AGCTCATCTTTTCAACGCAAGGATATGTTCTTTCTATGATGATCTGATGTCGATCACTA 1980
1921 AGCTCATCTTTTCAACGCAAGGATATGTTCTTTCTATGATGATCTGATGTCGATCACTA 1980
1981 GGGTACATGACAGCTGACGATTTCAAGACATTTCCCAAGATGATGATGATGATGATGATG 2040
1981 GGGTACATGACAGCTGACGATTTCAAGACATTTCCCAAGATGATGATGATGATGATGATG 2040
2041 AATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2100
2041 AATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2100
2101 GAAATGTTACGAGAGTTTGGATGGGCAATGCTGAAGATTTGATGATGATGATGATGATGATG 2160
2101 GAAATGTTACGAGAGTTTGGATGGGCAATGCTGAAGATTTGATGATGATGATGATGATGATG 2160
2161 GAAATGTTACGAGAGTTTGGATGGGCAATGCTGAAGATTTGATGATGATGATGATGATGATG 2220
2161 GAAATGTTACGAGAGTTTGGATGGGCAATGCTGAAGATTTGATGATGATGATGATGATGATG 2220
2221 GTGACAGTCTGACAGATGAAACCGGCTGCTGATGATGATGATGATGATGATGATGATGATG 2280
2221 GTGACAGTCTGACAGATGAAACCGGCTGCTGATGATGATGATGATGATGATGATGATGATG 2280
2281 TGTCTATCTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 2340
2281 TGTCTATCTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGAT 2340
2341 AGATACGAGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2400
2341 AGATACGAGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2400
2401 TCATTA 2406
2401 TCATTA 2406

RESULT 3
US-10-600-070-128
; Sequence 128, Application US/10600070
; Publication No. US20040139500A1
; GENERAL INFORMATION:
; APPLICANT: Ostryoung, Katherine W.
; APPLICANT: Vitna, Stanislaw
; APPLICANT: Kokeharova, Olga A.
; APPLICANT: Gao, Hongbo
; TITLE OF INVENTION: Plasmid Division and Related Genes and Proteins, and Methods of
; FILE REFERENCE: MSU-08153
; CURRENT APPLICATION NUMBER: US/10/600,070
; CURRENT FILING DATE: 2003-06-20
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 128
; LENGTH: 2406
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-600-070-128

Query Match 99.7%; Score 2399.6; DB 7; Length 2406;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2402; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
1 ATGGAAGCTCTGAGTCAAGTGGGCAATGGTCTCTCCCATTCGATTTATGCGGATTTACCA 60
1 ATGGAAGCTCTGAGTCAAGTGGGCAATGGTCTCTCTCCCATTCGATTTATGCGGATTTACCA 60

61 CCGGCGACGACAAAGCTCCGACGTAGGACACAAACCTCTAACATCTATCTGCTCCGACG 120
61 CCGGCGACGACAAAGCTCCGACGTAGGACACAAACCTCTAACATCTATCTGCTCCGACG 120
121 AATATGCGGACGACGCTCTCTCTCCGACTTCAATTTTCACTCCGATTTCTCTCTCTCTCT 180
121 AATATGCGGACGACGCTCTCTCTCCGACTTCAATTTTCACTCCGATTTCTCTCTCTCTCT 180
181 TTGCGCACCGGACACCAACCGCACTCTGCTCTCTGCGACCAATATGATGATGATGATGATG 240
181 TTGCGCACCGGACACCAACCGCACTCTGCTCTCTGCGACCAATATGATGATGATGATGATG 240
241 GAAAGCGACGCTCCCATTCGATTTGATTTCTACCAAGGATATGAGAGCTCAAAACATTTTC 300
241 GAAAGCGACGCTCCCATTCGATTTGATTTCTACCAAGGATATGAGAGCTCAAAACATTTTC 300
301 TTAAACCGATGGAATCAGAAAGCAATTTGAGGCTTTCCGAAACCGCGCAATTCGGT 360
301 TTAAACCGATGGAATCAGAAAGCAATTTGAGGCTTTCCGAAACCGCGCAATTCGGT 360
361 TTCAAGCAGACGCTTTAATCAGCCGAGACAGATTTCTCAAGCTGCTTCCGAAACTCTG 420
361 TTCAAGCAGACGCTTTAATCAGCCGAGACAGATTTCTCAAGCTGCTTCCGAAACTCTG 420
421 TCTAATCTCGGCTAGAGAGAGTACATGAAAGTCTTCTGATGATGATGATGATGATGATGATG 480
421 TCTAATCTCGGCTAGAGAGAGTACATGAAAGTCTTCTGATGATGATGATGATGATGATGATG 480
481 GTCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
481 GTCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
541 GGTAGACGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
541 GGTAGACGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600
601 TCGTTTAAAGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
601 TCGTTTAAAGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 660
661 GCTATGAGCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720
661 GCTATGAGCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 720
721 AAGCTTTTACAG 780
721 AAGCTTTTACAG 780
781 GAGACTTTGAG 840
781 GAGACTTTGAG 840
841 GATATATGCTGCGGAAAG 900
841 GATATATGCTGCGGAAAG 900
901 GTTGAAG 960
901 GTTGAAG 960
961 GAGGCGTTTTCAGATGACAGCTGCTGAGACAGGTTGATCTTTTGTAGCTACCCCAAGC 1020
961 GAGGCGTTTTCAGATGACAGCTGCTGAGACAGGTTGATCTTTTGTAGCTACCCCAAGC 1020
1021 AATATTTACAG 1080
1021 AATATTTACAG 1080
1081 TTTATGTTAG 1140
1081 TTTATGTTAG 1140

1141 CAGGCTAAGTAAATGCTATGAGATTCTCGAGTCTGTATGATACACGAAATATTG 1200
1141 CAGGCTAAGTAAATGCTATGAGATTCTCGAGTCTGTATGATACACGAAATATTG 1200
1201 GAGATAGACTTGCTGTAGAAAGGGGACCTGTGCACTGCTTATAGGCAAGTTGATG 1260
1201 GAGATAGACTTGCTGTAGAAAGGGGACCTGTGCACTGCTTATAGGCAAGTTGATG 1260
1261 TCCCGATATGCTGTGGCTTAAACAGTGAAGATTCATATATAGAAATCACTATTTG 1320
1261 TCCCGATATGCTGTGGCTTAAACAGTGAAGATTCATATATAGAAATCACTATTTG 1320
1321 GAGTTGTTTGGAGAAATTCAAATGCTATGACATGATGATCTCCCTGACATATGCAA 1380
1321 GAGTTGTTTGGAGAAATTCAAATGCTATGACATGATGATCTCCCTGACATATGCAA 1380
1381 TTGTTGAAACCTGGTGGCAGGGGTTGCTTTCCTAGGTTGAGAGACCAAGATTA 1440
1381 TTGTTGAAACCTGGTGGCAGGGGTTGCTTTCCTAGGTTGAGAGACCAAGATTA 1440
1441 AAATTTAACTCGGGGACTACTATGATGATCCATATGTTTGAATTAATTTGAAAGAG 1500
1441 AAATTTAACTCGGGGACTACTATGATGATCCATATGTTTGAATTAATTTGAAAGAG 1500
1501 GAGGTAGTTCAAGGTTCTCTCTTAACTGCTGCACTATGCGCAAGATTGAGCCGAG 1560
1501 GAGGTAGTTCAAGGTTCTCTCTTAACTGCTGCACTATGCGCAAGATTGAGCCGAG 1560
1561 CATGTGAAAGCTAGTGTATGAGGCACTGAGAAAGTTTTCCTCCCGCTATACAGT 1620
1561 CATGTGAAAGCTAGTGTATGAGGCACTGAGAAAGTTTTCCTCCCGCTATACAGT 1620
1621 AGAACTCGGCTGAACCCAGAGATGTGCAAGACAGTGTATGATGATCTGTTGGT 1680
1621 AGAACTCGGCTGAACCCAGAGATGTGCAAGACAGTGTATGATGATCTGTTGGT 1680
1681 AACAATGTAAGCCGTAATGCTGAGCTGCTGCTTATATGCAAGAGCTGTAAACCTCT 1740
1681 AACAATGTAAGCCGTAATGCTGAGCTGCTGCTTATATGCAAGAGCTGTAAACCTCT 1740
1741 GAAACTTTGAAACTAATGATTAATGCAATTCAGAGCTGGGCTTCAGAGATGAGCTTGA 1800
1741 GAAACTTTGAAACTAATGATTAATGCAATTCAGAGCTGGGCTTCAGAGATGAGCTTGA 1800
1801 GAAACTTTGAAACTAATGATTAATGCAATTCAGAGCTGGGCTTCAGAGATGAGCTTGA 1860
1801 GAAACTTTGAAACTAATGATTAATGCAATTCAGAGCTGGGCTTCAGAGATGAGCTTGA 1860
1861 GCTGCTGCTGTGGCAATGGAATGCTGATTAATGTTAAAGAGGCAAGTGTGAAGTCTTA 1920
1861 GCTGCTGCTGTGGCAATGGAATGCTGATTAATGTTAAAGAGGCAAGTGTGAAGTCTTA 1920
1921 AGCTATCTTTTCAACGAGAGATGATGTTCTTATAGGAATCTGATGCTGCTTCACTA 1980
1921 AGCTATCTTTTCAACGAGAGATGATGTTCTTATAGGAATCTGATGCTGCTTCACTA 1980
1981 GGGTCAGTCAAGCTGAGATTCAGAAAGCACTTCCAGAAATGATGATGATGATGATG 2040
1981 GGGTCAGTCAAGCTGAGATTCAGAAAGCACTTCCAGAAATGATGATGATGATGATG 2040
1991 GGGTCAGTCAAGCTGAGATTCAGAAAGCACTTCCAGAAATGATGATGATGATGATG 2040
1991 GGGTCAGTCAAGCTGAGATTCAGAAAGCACTTCCAGAAATGATGATGATGATGATG 2040
2041 AATATAGTATCCAGAGATTCAGAAAGCACTTCCAGAAATGATGATGATGATGATG 2100
2041 AATATAGTATCCAGAGATTCAGAAAGCACTTCCAGAAATGATGATGATGATGATG 2100
2101 GAAATGTTTACCAAGAGATTCAGAAAGCACTTCCAGAAATGATGATGATGATGATG 2160
2101 GAAATGTTTACCAAGAGATTCAGAAAGCACTTCCAGAAATGATGATGATGATGATG 2160
2161 GAAATGTTTACCAAGAGATTCAGAAAGCACTTCCAGAAATGATGATGATGATGATG 2220
2161 GAAATGTTTACCAAGAGATTCAGAAAGCACTTCCAGAAATGATGATGATGATGATG 2220
2221 GTGACAGTCTCAGCAGATGAAACCCGCTGCTGCTGAGAAACCACTTGAAGAGTCTGCT 2280

2221 GTGACAGTCTCAGCAGATGAAACCCGCTGCTGCTGAGAAACCACTTGAAGAGTCTGCT 2280
2281 TGTATCTGATTTGGTTCATCCAGAAACCAATGCTATGATGATGATGATGATGATG 2340
2281 TGTATCTGATTTGGTTCATCCAGAAACCAATGCTATGATGATGATGATGATGATG 2340
2341 AGATACGAAGTTTCTGCTGCTCAAGTCAAGGTGGAAGAAATCACTGAAGCTCTGTTGCA 2400
2341 AGATACGAAGTTTCTGCTGCTCAAGTCAAGGTGGAAGAAATCACTGAAGCTCTGTTGCA 2400
2401 TCATTA 2406
2401 TCATTA 2406

RESULT 4
US-10-600-070-130
; Sequence 130, Application US/10600070
; Publication No. US20040139500A1
; GENERAL INFORMATION:
; APPLICANT: Oseeryoung, Katherine W.
; APPLICANT: Vittha, Stanislaw
; APPLICANT: Koksharova, Olga A.
; APPLICANT: Gao, Hong
; TITLE OF INVENTION: Plasmid Division and Related Genes and Proteins, and Methods of
; FILE REFERENCE: MSU-08153
; CURRENT APPLICATION NUMBER: US/10/600,070
; CURRENT FILING DATE: 2003-06-20
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 130
; LENGTH: 2637
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-600-070-130

Query Match 99.7%; Score 2399.6; DB 7; Length 2637;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2402; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

1 ATGGAAGCTCTGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 60
114 ATGGAAGCTCTGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 173
61 CCGGCGAGAGAAAGCTCCGAGCTGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 120
174 CCGGCGAGAGAAAGCTCCGAGCTGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 233
121 AATGAGGCGAGCGCTCTCTCCGACTTCAATTTCACTCCGATTCCTCCCTCCCTCC 180
234 AATGAGGCGAGCGCTCTCTCCGACTTCAATTTCACTCCGATTCCTCCCTCCCTCC 293
181 TTGCGCAACCGGCAACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 240
294 TTGCGCAACCGGCAACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 353
241 GAAAGCCAGCTCCGATCCGATGATTTTCAAGAGATTTGAGAGCTCAAGCAAGCTTTC 300
354 GAAAGCCAGCTCCGATCCGATGATTTTCAAGAGATTTGAGAGCTCAAGCAAGCTTTC 413
301 TTAACCGATGATCAAGAGATTTGAGAGATTTGAGAGATTTGAGAGATTTGAGAGATTT 360
414 TTAACCGATGATCAAGAGATTTGAGAGATTTGAGAGATTTGAGAGATTTGAGAGATTT 473
361 TTACGCGAGAGCTTTTATATCAGCGAGAGAGATTTTCAAGAGCTCTTGGAAGCTCTG 420
474 TTACGCGAGAGCTTTTATATCAGCGAGAGAGATTTTCAAGAGCTCTTGGAAGCTCTG 533
421 TCTATATCTCGGTCTTAAAGAGATTAAGAGATTTTCAAGAGCTCTTGGAAGCTCTG 480
534 TCTATATCTCGGTCTTAAAGAGATTAAGAGATTTTCAAGAGCTCTTGGAAGCTCTG 593

QY 481 GTCACTACTGATGTTCTTGGGATAGGTTCTTGGGAGCTCTCTGTGTAATGGCAAGAGT 540
Db 594 GTCACTACTGATGTTCTTGGGATAGGTTCTTGGGAGCTCTCTGTGTAATGGCAAGAGT 653
QY 541 GGTGAGACTGAGATGTTCTTGGGATAGGTTCTTGGGAGCTCTCTGTGTAATGGCAAGAGT 600
Db 654 GGTGAGACTGAGATGTTCTTGGGATAGGTTCTTGGGAGCTCTCTGTGTAATGGCAAGAGT 713
QY 601 TCGTTTAAGCAAGATGTTGTTTAAATGTTAGGCGCTTGGGATGTTCTGATGTTCTGAGGAT 660
Db 714 TCGTTTAAGCAAGATGTTGTTTAAATGTTAGGCGCTTGGGATGTTCTGATGTTCTGAGGAT 773
QY 661 GCTATGCGATGATGATCACTGATTTTAACTGTGTTAATGATTTGTTGAGGAACTTTG 720
Db 774 GCTATGCGATGATGATCACTGATTTTAACTGTGTTAATGATTTGTTGAGGAACTTTG 833
QY 721 AAGCTTTTACAGAGAGAGAGCAAGTACCTTGGACCGGATTTACGTGCAAAATTTAT 780
Db 834 AAGCTTTTACAGAGAGAGAGCAAGTACCTTGGACCGGATTTACGTGCAAAATTTAT 893
QY 781 GAGACTTTGAAAGAGATCACTCCGCTTATGTTGAGGACTACTTGGCTTACCGCTTGGT 840
Db 894 GAGACTTTGAAAGAGATCACTCCGCTTATGTTGAGGACTACTTGGCTTACCGCTTGGT 953
QY 841 GATGATTAACGTCGCAAAAGACTAAATGTTTAAAGCGGTGTGCGGAATATTTTGTGCT 900
Db 954 GATGATTAACGTCGCAAAAGACTAAATGTTTAAAGCGGTGTGCGGAATATTTTGTGCT 1013
QY 901 GTTGAAGAGAGTGAAGCATCAGCTCTTGTGGGGTTTGAACCGGATTAATTTATGAT 960
Db 1014 GTTGAAGAGAGTGAAGCATCAGCTCTTGTGGGGTTTGAACCGGATTAATTTATGAT 1073
QY 961 GAGCGTTTTCAGATGACAGCTGCTGAGAGGTTGATCTTTTGTGATACCCCAAGC 1020
Db 1074 GAGCGTTTTCAGATGACAGCTGCTGAGAGGTTGATCTTTTGTGATACCCCAAGC 1133
QY 1021 AATATTCAGAGAGTCAATTTGAAGTTTACGAAGTTGCACTTGTCTTGTGGCTCAAGCT 1080
Db 1134 AATATTCAGAGAGTCAATTTGAAGTTTACGAAGTTGCACTTGTCTTGTGGCTCAAGCT 1193
QY 1081 TTTATTTGTAAGAAAGCCACCTTTTACAGATGCTGATTAAGCAATTCACAGACTTATG 1140
Db 1194 TTTATTTGTAAGAAAGCCACCTTTTACAGATGCTGATTAAGCAATTCACAGACTTATG 1253
QY 1141 CAGGCTAAGTATAGGCTATGAGATTCCTGCGATGTTGATGATACAGGAAATTTGG 1200
Db 1254 CAGGCTAAGTATAGGCTATGAGATTCCTGCGATGTTGATGATACAGGAAATTTGG 1313
QY 1201 GAGATTAAGCTTCGCTCTAGAAAGGGAATCTGTGCACTGCTTATAGGCAAAAGTTGATGA 1260
Db 1314 GAGATTAAGCTTCGCTCTAGAAAGGGAATCTGTGCACTGCTTATAGGCAAAAGTTGATGA 1373
QY 1261 TGGCGTAAATGTTGGGCTTGAAGATGAGATTCACAAATATAGGAATTCAGCTATTTG 1320
Db 1374 TGGCGTAAATGTTGGGCTTGAAGATGAGATTCACAAATATAGGAATTCAGCTATTTG 1433
QY 1321 GAGTTTGTGTTGAGAAATTCAAATCGTATGATGACAAATGATGATCTCCCTGAGCTATGCAA 1380
Db 1434 GAGTTTGTGTTGAGAAATTCAAATCGTATGATGACAAATGATGATCTCCCTGAGCTATGCAA 1493
QY 1381 TTGTTGAAACCTGTTGGAGGGGTTGCTTTCTTGAAGTTCAAGACCAAAAGATTA 1440
Db 1494 TTGTTGAAACCTGTTGGAGGGGTTGCTTTCTTGAAGTTCAAGACCAAAAGATTA 1553
QY 1441 AAATTTAACTCGGGGACTACTATGATATGCTTATGTTGAGTTACTTGGAAAGAGTG 1500
Db 1554 AAATTTAACTCGGGGACTACTATGATATGCTTATGTTGAGTTACTTGGAAAGAGTG 1613
QY 1501 GAGGTAGTTGAGGGTTCTCTTTAGACTCTCTGCAATATGAGCAAGATTTGAGCCGAG 1560
Db 1614 GAGGTAGTTGAGGGTTCTCTTTAGACTCTCTGCAATATGAGCAAGATTTGAGCCGAG 1673

QY 1561 CATGTGAAGCTAGTGTATGACAGCACTGACAGAAATTTTCTTCCCTGATATACAT 1620
Db 1674 CATGTGAAGCTAGTGTATGACAGCACTGACAGAAATTTTCTTCCCTGATATACAT 1733
QY 1621 AGAACTCGGCTGAACCCAGAGATGTCAGAGACATGTTTATGATGATCCCTTGGT 1680
Db 1734 AGAACTCGGCTGAACCCAGAGATGTCAGAGACATGTTTATGATGATCCCTTGGT 1793
QY 1681 AACATGATGAGCCCTGATGAGCTGAGCTGTTCTTTATGCAAGAGCTGTAAAGCTCT 1740
Db 1794 AACATGATGAGCCCTGATGAGCTGAGCTGTTCTTTATGCAAGAGCTGTAAAGCTCT 1853
QY 1741 GAAACTTTTGAACCTAATGATTAATGCAATTCAGCTGAGGATCTCAGAGATGAGCTTAT 1800
Db 1854 GAAACTTTTGAACCTAATGATTAATGCAATTCAGCTGAGGATCTCAGAGATGAGCTTAT 1913
QY 1801 GAACTACTGTTGAAATGTCCTGTTGCTGATATGTTAAAGAGGCAAGTGTGAATCTTA 1860
Db 1914 GAACTACTGTTGAAATGTCCTGTTGCTGATATGTTAAAGAGGCAAGTGTGAATCTTA 1973
QY 1861 GCTGCTGTTGAGCAATTTGCACTGATTTCACTGTTCAAGCCAGAGATATTTTCTTAAAGC 1920
Db 1974 GCTGCTGTTGAGCAATTTGCACTGATTTCACTGTTCAAGCCAGAGATATTTTCTTAAAGC 2033
QY 1921 AGCTCATCTTTTCAAGCCAGAGATGAGTATGTTTCTATGAAATCTGATGCTTACATA 1980
Db 2034 AGCTCATCTTTTCAAGCCAGAGATGAGTATGTTTCTATGAAATCTGATGCTTACATA 2093
QY 1981 GGGTCAGTCAGAGCTGACGATTCAGAGCACTTCCAGATGATGATGCTAGACCTGACAG 2040
Db 2094 GGGTCAGTCAGAGCTGACGATTCAGAGCACTTCCAGATGATGATGCTAGACCTGACAG 2153
QY 2041 AATATGATTCAGAGTGGCAGAGATTAAGTCTCTGCTTTTGGGCTGATCAACCGATA 2100
Db 2154 AATATGATTCAGAGTGGCAGAGATTAAGTCTCTGCTTTTGGGCTGATCAACCGATA 2213
QY 2101 GAAATGTTTACAGAGTTTGGATGGCGAATGCTGAAGATTTGAGACTGACAGAGACT 2160
Db 2214 GAAATGTTTACAGAGTTTGGATGGCGAATGCTGAAGATTTGAGACTGACAGAGACT 2273
QY 2161 GAACTGCGAGCTGTTGGTTTATGATTAATCACTGTGTAACATGCTTGTGAAGT 2220
Db 2274 GAACTGCGAGCTGTTGGTTTATGATTAATCACTGTGTAACATGCTTGTGAAGT 2333
QY 2221 GTGACAGTCTCAGAGATGGAACCCGCTCTGCTGAGAACTCTGAGAGAGTCTGCT 2280
Db 2334 GTGACAGTCTCAGAGATGGAACCCGCTCTGCTGAGAACTCTGAGAGAGTCTGCT 2393
QY 2281 TGCTATCTGATTTGTTCAATCCAGAAACATGCTATGATGTCAGAACTTACACACA 2340
Db 2394 TGCTATCTGATTTGTTCAATCCAGAAACATGCTATGATGTCAGAACTTACACACA 2453
QY 2341 AGATACGAAAGTTTCTGCTCAGAGTCAAGGATGAGAAATCACTGAAGGCTCTGTTCTTGA 2400
Db 2454 AGATACGAAAGTTTCTGCTCAGAGTCAAGGATGAGAAATCACTGAAGGCTCTGTTCTTGA 2513
QY 2401 TCATTA 2406
Db 2514 TCATTA 2519

RESULT 5
US-10-739-930-227
; Sequence 227, Application US/10739930
; Publication No. US20040216190A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; FILE REFERENCE: 38-21(53377)B
; CURRENT APPLICATION NUMBER: US/10/739, 930
; CURRENT FILING DATE: 2003-12-18
; NUMBER OF SEQ ID NOS: 11088

```
; SEQ ID NO 227
; LENGTH: 2679
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: Clone ID: ARATH-23APR03-CLUSTER13643_1
; US-10-739-930-227

Query Match      99.7%; Score 2399.6; DB 8; Length 2679;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2402; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGGAAGCTCTGAGTCACTGCGCATTTGCTCTCCCATTCCTCAATTAATGCGGATTACCA 60
DB 115 ATGGAAGCTCTGAGTCACTGCGCATTTGCTCTCCCATTCCTCAATTAATGCGGATTACCA 174
QY 61 CCGGGAGACGACAAAGCTCCGAGTAGCCCAACACTCTACAACTATGCTCCGACAG 120
DB 175 CCGGGAGACGACAAAGCTCCGAGTAGCCCAACACTCTACAACTATGCTCCGACAG 234
QY 121 AATAGGCGCGACCGCTCTCTCCGACTTCAATTTCACTCCGATTCCTCTCTCTCC 180
DB 235 AATAGGCGCGACCGCTCTCTCTCCGACTTCAATTTCACTCCGATTCCTCTCTCTCC 294
QY 181 TTGCGCACCGCCACCAACCGCCCACTCTGCTCTCTGCGACCAATTAATGATCGTCCC 240
DB 295 TTGCGCACCGCCACCAACCGCCCACTCTGCTCTCTGCGACCAATTAATGATCGTCCC 354
QY 241 GAAGCGCAGTCCCATCCCATTTGATTTCTACAGAGTATTAAGAGCTTCAACATTTTC 300
DB 355 GAAGCGCAGTCCCATCCCATTTGATTTCTACAGAGTATTAAGAGCTTCAACATTTTC 414
QY 301 TTAAACGATGGAATGAGAAGACATTCGAAAGCTAGGGTTTCGAAACCGCGCATTCGGT 360
DB 415 TTAAACGATGGAATGAGAAGACATTCGAAAGCTAGGGTTTCGAAACCGCGCATTCGGT 474
QY 361 TTACAGGAGACGCTTTTAATCAAGCGGAGACAGATTTCTTAAGCTGCTTTCGAAACTCTG 420
DB 475 TTACAGGAGACGCTTTTAATCAAGCGGAGACAGATTTCTTAAGCTGCTTTCGAAACTCTG 534
QY 421 TCTAATCCCTCGGTCTAGAAAGAGTACAAATGAAGGCTCTTGATGATGAAGAAACTTCA 480
DB 535 TCTAATCCCTCGGTCTAGAAAGAGTACAAATGAAGGCTCTTGATGATGAAGAAACTTCA 594
QY 481 GTCATCACTGATGTTCTTGGGATTAAGTCTCTGGGGCTCTCTGATTAATGCAAGAAGT 540
DB 595 GTCATCACTGATGTTCTTGGGATTAAGTCTCTGGGGCTCTCTGATTAATGCAAGAAGT 654
QY 541 GGTGAGACTGAGATGTTCTTGGGTTGGTGAAGGCTCTGCTTAAGAGAAGTTCCTAAG 600
DB 655 GGTGAGACTGAGATGTTCTTGGGTTGGTGAAGGCTCTGCTTAAGAGAAGTTCCTAAG 714
QY 601 TCGTTAAGCAAGATGTTTAACTTAATGCGGCTTGGCTTCTGATGATCTCCGAGGAT 660
DB 715 TCGTTAAGCAAGATGTTTAACTTAATGCGGCTTGGCTTCTGATGATCTCCGAGGAT 774
QY 661 GCTATGCACTGATTCACACTGATTTTAACTGCTTAAGATTTGTTGAAGAAGCTTTG 720
DB 775 GCTATGCACTGATTCACACTGATTTTAACTGCTTAAGATTTGTTGAAGAAGCTTTG 834
QY 721 AAGCTTTTACAGAGGAAGAGCAAGATAGCTTGCACCGGATTTACGTCACAAATTAAT 780
DB 835 AAGCTTTTACAGAGGAAGAGCAAGATAGCTTGCACCGGATTTACGTCACAAATTAAT 894
QY 781 GAGACTTTGGAAGAGATCACTCCGGGTATGCTTGGAGCTACTTGGCTTACCGGTTGGT 840
DB 895 GAGACTTTGGAAGAGATCACTCCGGGTATGCTTGGAGCTACTTGGCTTACCGGTTGGT 954
QY 841 GATGATTAACGCTGCGAAAAGACTAAATGATTTAAGCGGTGTCGGAATATTTTGTGTCT 900
DB 955 GATGATTAACGCTGCGAAAAGACTAAATGATTTAAGCGGTGTCGGAATATTTTGTGTCT 1014
QY 901 GTTGAAGAGGTGAGGATCACTCTTGTGTGGGGGTTTGAACCGGTGAGAAAGTTATGAT 960
```

```
DB 1015 GTTGAAGAGGTGAGGATCACTCTTGTGGGGGTTTGAACCGGTGAGAAAGTTATGAT 1074
QY 961 GAGGCGTTTTTACGAATACAGCTGCTGAGACAGGTGATCTTTTGTAGCTAACCCCAAGC 1020
DB 1075 GAGGCGTTTTTACGAATACAGCTGCTGAGACAGGTGATCTTTTGTAGCTAACCCCAAGC 1134
QY 1021 AATATTCAGACAGATCAATTTGAAGTTTACGAAGTTGACATTCCTTGTGGCTCAAGCT 1080
DB 1135 AATATTCAGACAGATCAATTTGAAGTTTACGAAGTTGACATTCCTTGTGGCTCAAGCT 1194
QY 1081 TTTATTTGTAAGAAAGCCACACCTTTTACAGATGCTGATTAAGCAATTCACGACATTCAG 1140
DB 1195 TTTATTTGTAAGAAAGCCACACCTTTTACAGATGCTGATTAAGCAATTCACGACATTCAG 1254
QY 1141 CAGGCTAAGTAAATGAGCTAATGAGATTCCTGAGATGTTGATTAATGATACCGGAATATG 1200
DB 1255 CAGGCTAAGTAAATGAGCTAATGAGATTCCTGAGATGTTGATTAATGATACCGGAATATG 1314
QY 1201 GAGATAGACTTCGGCTTAGAAAGGGGACTCTGTGCACTGCTTATAGGCAAGTTGATGAA 1260
DB 1315 GAGATAGACTTCGGCTTAGAAAGGGGACTCTGTGCACTGCTTATAGGCAAGTTGATGAA 1374
QY 1261 TTCCGATATGTTGGGCTTAGAACAATGAGGATTCACAAATTAAGGAATCCAGCTATTGTG 1320
DB 1375 TTCCGATATGTTGGGCTTAGAACAATGAGGATTCACAAATTAAGGAATCCAGCTATTGTG 1434
QY 1321 GAGTTGTTTTTGGGAATTCAAATCGATGATGACAAATGATATCCCTCGACATATGCAAA 1380
DB 1435 GAGTTGTTTTTGGGAATTCAAATCGATGATGACAAATGATATCCCTCGACATATGCAAA 1494
QY 1381 TTGTTGGAACCTGTTGGCAAGGGGTTGCTTCTCTAGGTTCAAGAGACACCAAGATTA 1440
DB 1495 TTGTTGGAACCTGTTGGCAAGGGGTTGCTTCTCTAGGTTCAAGAGACACCAAGATTA 1554
QY 1441 AATTTAAACTCGGGGACTACTATGATGATCTTAATGTTTGAATTAATTGAAAAGATG 1500
DB 1555 AATTTAAACTCGGGGACTACTATGATGATCTTAATGTTTGAATTAATTGAAAAGATG 1614
QY 1501 GAGGTAGTCAAGGTTTCCTTTAGCTGCTGCTGCAACTATGGAAGATTGGAGCCGAG 1560
DB 1615 GAGGTAGTCAAGGTTTCCTTTAGCTGCTGCTGCAACTATGGAAGATTGGAGCCGAG 1674
QY 1561 CATGTGAAGCTAATGCTATGATGACAGCACTGCAAGAAAGTTTCTTCCCTCCGCTATACAGAT 1620
DB 1675 CATGTGAAGCTAATGCTATGATGACAGCACTGCAAGAAAGTTTCTTCCCTCCGCTATACAGAT 1734
QY 1621 AGAAACTCGGCTGAACCCCAAGGATGTGCAAGAGACAGTGTTAATGATCTGTTGGT 1680
DB 1735 AGAAACTCGGCTGAACCCCAAGGATGTGCAAGAGACAGTGTTAATGATCTGTTGGT 1794
QY 1681 AACAAATGAGCGCGTGAATGATGAGACCTGCTGCTTATGAGAGAAAGCTGTAAGACCCCT 1740
DB 1795 AACAAATGAGCGCGTGAATGATGAGACCTGCTGCTTATGAGAGAAAGCTGTAAGACCCCT 1854
QY 1741 GAAAACTTTGAAACTAATGATTAATGCAATTCGACTGCGGGCTCAGAGAGTAGGTTGAT 1800
DB 1855 GAAAACTTTGAAACTAATGATTAATGCAATTCGACTGCGGGCTCAGAGAGTAGGTTGAT 1914
QY 1801 GAAACTACTGTTGAATATGTCGGTGTGATATGTTAAAGAGAGCAAGTGTGAAGATCCTA 1860
DB 1915 GAAACTACTGTTGAATATGTCGGTGTGATATGTTAAAGAGAGCAAGTGTGAAGATCCTA 1974
QY 1861 GCTGCTGTTGGCAATTTGACATGATTTCACTGTTACGCGACGAAAGTATTTTCTTAAAGC 1920
DB 1975 GCTGCTGTTGGCAATTTGACATGATTTCACTGTTACGCGACGAAAGTATTTTCTTAAAGC 2034
QY 1921 AGCTCATCTTTTCAACCAAGATATGATTTCTTCTATGGAATCTGATGCTGATACATA 1980
DB 2035 AGCTCATCTTTTCAACCAAGATATGATTTCTTCTATGGAATCTGATGCTGATACATA 2094
QY 1981 GGGTCAGTCAAGCTGACGATTCAGAAAGCACTTCCCAATGATGATGCTAGACTGACAG 2040
```

Db 2095 GGGTCAGTCAGAGTCAGATTGAGAACCTTCCAGAAATGATGCTAGAGCTGACAG 2154
Qy 2041 AATAATGATCCAGTGGCAGAAAGATTAACTCTGCTTTTGGGCTGATCACCGCAT 2100
Db 2155 AATATAGATCCAGTGGCAGAAAGATTAACTCTGCTTTTGGGCTGATCACCGCAT 2214
Qy 2101 GAAATGTTACAGAGTTTGGATGGGCAATGCTGAAATTTGGACTGACAGAGCAGT 2160
Db 2215 GAAATGTTACAGAGTTTGGATGGGCAATGCTGAAATTTGGACTGACAGAGCAGT 2274
Qy 2161 GAACTGGGAGCTTTGGGTTTATGATTAATACATGTTGAAATCTATCTGTTGACAGT 2220
Db 2275 GAACTGGGAGCTTTGGGTTTATGATTAATACATGTTGAAATCTATCTGTTGACAGT 2334
Qy 2221 GTGACAGTCTGACAGATGAAACCCGCTCTGCTGTTGAAAGCACTCTGGAAGTCTCT 2280
Db 2235 GTGACAGTCTGACAGATGAAACCCGCTCTGCTGTTGAAAGCACTCTGGAAGTCTCT 2334
Qy 2281 TGTCTATCTGATTTGGTTCATTCAGAAAACAATGCTACTGATGTCAGAACTTACACACA 2340
Db 2395 TGTCTATCTGATTTGGTTCATTCAGAAAACAATGCTACTGATGTCAGAACTTACACACA 2454
Qy 2341 AGATACGAGTTTCTGCTCCAGTCAAGGTGAAAAATCACTGAAGCTCTGTTCTTGA 2400
Db 2455 AGATACGAGTTTCTGCTCCAGTCAAGGTGAAAAATCACTGAAGCTCTGTTCTTGA 2514
Qy 2401 TCATTA 2406
Db 2515 TCATTA 2520

RESULT 6
us-10-600-070-3
: Sequence 3, Application US/10600070
: Publication No. US20040139500A1
: GENERAL INFORMATION:
: APPLICANT: Osteoryoung, Katherine W.
: APPLICANT: Vittha, Stanislaw
: APPLICANT: Kokeharova, Olga A.
: APPLICANT: Gao, Hongbo
: TITLE OF INVENTION: Placid Division and Related Genes and Proteins, and Methods of
: TITLE OF INVENTION: Use
: FILE REFERENCE: MSU-08153
: CURRENT APPLICATION NUMBER: US/10/600,070
: CURRENT FILING DATE: 2003-06-20
: NUMBER OF SEQ ID NOS: 206
: SOFTWARE: PatentIn version 3.2
: SEQ ID NO 3
: LENGTH: 3667
: TYPE: DNA
: ORGANISM: Arabidopsis thaliana
US-10-600-070-3

Query Match 71.7%; Score 1724; DB 7; Length 3667;
Best Local Similarity 79.2%; Pred. No. 0;
Matches 2406; Conservative 0; Mismatches 0; Indels 632; Gaps 5;

Qy 1 ATGGAAGCTGAGTCACTGCGGATGATGATCTCCCATTCGAATTATGCGATTACCA 60
Db 481 ATGGAAGCTGAGTCACTGCGGATGATGATGATCTCCCATTCGAATTATGCGATTACCA 540
Qy 61 CCGGAGCAGAAAGCTCCGAGTACGACAAACACTTATGATCTGCTCCGAGC 120
Db 541 CCGGAGCAGAAAGCTCCGAGTACGACAAACACTTATGATCTGCTCCGAGC 600
Qy 121 AAATGGGCGGACCGTCTCTCTCCGACCTTCAATTTCACTCGAGTCTCTCTCTCC 180
Db 601 AAATGGGCGGACCGTCTCTCTCCGACCTTCAATTTCACTCGAGTCTCTCTCTCC 660
Qy 181 TTGGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 240
Db 661 TTGGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 720

Qy 241 GAACGCCAGTCCCATCCCATTTGATTTCTACAGATATTAAGAGCTCAACAATTC 300
Db 721 GAACGCCAGTCCCATCCCATTTGATTTCTACAGATATTAAGAGCTCAACAATTC 780
Qy 301 TTAAACCATGGAATTCAGAAAGAGATTCGAGCTTAAGGTTTCGAAACCGCGCAATTCGT 360
Db 781 TTAAACCATGGAATTCAGAAAGAGATTCGAGCTTAAGGTTTCGAAACCGCGCAATTCGT 840
Qy 361 TTGACGACAGCAGCTTTAATCAGCCGAGACAGATTTCTCAAGCTGCTCGAAATCTCTG 420
Db 841 TTGACGACAGCAGCTTTAATCAGCCGAGACAGATTTCTCAAGCTGCTCGAAATCTCTG 900
Qy 421 TCTAATCTCGGCTTGAAGAGATTAAGATTAAGAGTCTTCTGATGATGAAGAGCTTACA 480
Db 901 TCTAATCTCGGCTTGAAGAGATTAAGATTAAGAGTCTTCTGATGATGAAGAGCTTACA 960
Qy 481 GTCAATCACTGATGTTCTTGGGAT-----AGGTTCTGAGGAT 504
Db 961 GTCAATCACTGATGTTCTTGGGAT-----AGGTTCTGAGGAT 1020
Qy 505 -----AGGTTCTGAGGAT 518
Db 1021 GTTTAATTTCAATGATTAAGAGAGAACTTTATCTAGTGAAGTTCTTGGGAT 1080
Qy 519 TCTCTGATATTCAGAGAGTGTGAGACTGAGATGTTCTTGGGATGAGCTCT 578
Db 1081 TCTCTGATATTCAGAGAGTGTGAGACTGAGATGTTCTTGGGATGAGCTCT 1140
Qy 579 GCTTAAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 638
Db 1141 GCTTAAGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 1200
Qy 639 GTTTCGATGATTCGAGGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 698
Db 1201 GTTTCGATGATTCGAGGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 1260
Qy 699 TGAATTTGTTGAGAGCTTTGAAGCTTTTAC----- 730
Db 1261 TGAATTTGTTGAGAGCTTTGAAGCTTTTAC----- 1320
Qy 731 -----A 731
Db 1321 ACGAGCTTGGCTTATTAAGAACTTTTATGATTTGATTTGATTTGATTTGATTTGAT 1380
Qy 732 GAGAGAGAGAGCAAGTGTGAGCTTGTGACCGGATTTACGTGACAAATGATGAGCTTTGA 791
Db 1381 GAGAGAGAGAGCAAGTGTGAGCTTGTGACCGGATTTACGTGACAAATGATGAGCTTTGA 1440
Qy 792 AGAGATCACTCCGCTTATGTTGAGAGTACTTGGCTTACCGCTTGGTATGATTAAGC 851
Db 1441 AGAGATCACTCCGCTTATGTTGAGAGTACTTGGCTTACCGCTTGGTATGATTAAGC 1500
Qy 852 TGGAAAAAGCTAAATGTTTAAAGCGGTGCGGAATTTTGGTCTGTTGAGAGAG 911
Db 1501 TGGAAAAAGCTAAATGTTTAAAGCGGTGCGGAATTTTGGTCTGTTGAGAGAG 1560
Qy 912 TGGAGCATGAGCTTCTGTTGGGTTTGAACCGGTGAGAACTTTATGATGAGAGGCTTTT 971
Db 1561 TGGAGCATGAGCTTCTGTTGGGTTTGAACCGGTGAGAACTTTATGATGAGAGGCTTTT 1620
Qy 972 ACGAATGACAGCTGCTAGC----- 991
Db 1621 ACGAATGACAGCTGCTAGC----- 1680
Qy 992 -----AGGTTGATCTTTTGT 1008
Db 1681 TGAATTAATCTTAAGTTTCTCATTTTATGATGATTTGATGATTTGATGATTTGAT 1740
Qy 1009 GCTACCCCAAGCAATATTCAGAGAGTCAATTTGAATTTGAAGTGTGACCTTGTCTT 1068
Db 1741 GCTACCCCAAGCAATATTCAGAGAGTCAATTTGAATTTGAAGTGTGACCTTGTCTT 1800
Qy 1069 GTGGCTCAAGCTTTTATGTTGATGAAGCAACCTTTTACAGAGTGTGATTAAGCAATTC 1128

```

Db      1801 GGGGCTCAAGCTTTATGTGTAAGAACCACTTTACAGAGATGCTGATAGCAATTC 1860
Qy      1129 GAGCACTTCAGAGGCTAAGGTAATGCTATGAGATTCGCGCATGTGTGATATCA 1188
Db      1861 CAGCACTTCAGAGGCTAAGGTAATGCTATGAGATTCGCGCATGTGTGATATCA 1920
Qy      1189 CGGAATTAATGGAGATAGACTTCGGTCTAGAAAGGGGACTCTGTGCACTGCTTAAGGC 1248
Db      1921 CGGAATTAATGGAGATAGACTTCGGTCTAGAAAGGGGACTCTGTGCACTGCTTAAGGC 1980
Qy      1249 AAAGTTGATGATGCCGTAATGTGGTTGGGCTTAGACAGTGAAGATTCACAAATATAGAAAT 1308
Db      1981 AAAGTTGATGATGCCGTAATGTGGTTGGGCTTAGACAGTGAAGATTCACAAATATAGAAAT 2040
Qy      1309 CCAGCTATGTTGAGAGTTGTTTGGAGATTCAAATCCGATGACAAATGATATCCCT 1368
Db      2041 CCAGCTATGTTGAGAGTTGTTTGGAGATTCAAATCCGATGACAAATGATATCCCT 2100
Qy      1369 GGACTATGCAAAATTTGTGAAAACCTGTGGCAGGGGTTGCTTCTTAGGTTCAAGAC 1428
Db      2101 GGACTATGCAAAATTTGTGAAAACCTGTGGCAGGGGTTGCTTCTTAGGTTCAAGAC 2160
Qy      1429 ACCAAGATTAATAATTTTAACTCGGGGACTATATGATGATCCTATGTTGTTGTTAC 1488
Db      2161 ACCAAGATTAATAATTTTAACTCGGGGACTATATGATGATCCTATGTTGTTGTTAC 2220
Qy      1489 TTGAAAAGTGGAGGTGTTGAGGGTTCTCCTTTAGCTGCTGCAACTATAGCAAGG 1548
Db      2221 TTGAAAAGTGGAGGTGTTGAGGGTTCTCCTTTAGCTGCTGCAACTATAGCAAGG 2280
Qy      1549 ATTGAGCCGAGCAGTGAAGAGCTAGTGTATGAGCAGCAGTGCAGAAAGTTTCTTCC 1608
Db      2281 ATTGAGCCGAGCAGTGAAGAGCTAGTGTATGAGCAGCAGTGCAGAAAGTTTCTTCC 2340
Qy      1609 CGCTATACAGATAGAACTCGGCTGAACCCAGAGATGTGCAAGAGACAGTGTTAAGTGA 1668
Db      2341 CGCTATACAGATAGAACTCGGCTGAACCCAGAGATGTGCAAGAGACAGTGTTAAGTGA 2400
Qy      1669 GATCGTGTGTTGTTAATGATGAGGCGGTGATGTGTGAGCCGTGCTGCTTATAGCAAGCT 1728
Db      2401 GATCGTGTGTTGTTAATGATGAGGCGGTGATGTGTGAGCCGTGCTGCTTATAGCAAGCT 2460
Qy      1729 GTAAGACCCCTCTGAAGAACTTTGAAACTATATGATATGCAATTCGAGCTGGGCTCTCAGAG 1788
Db      2461 GTAAGACCCCTCTGAAGAACTTTGAAACTATATGATATGCAATTCGAGCTGGGCTCTCAGAG 2520
Qy      1789 AGTACGCTTGTGATGAACCTACTGTTGAAATGTCCTGCTGATATGTTAAAGAGGCAAGT 1848
Db      2521 AGTACGCTTGTGATGAACCTACTGTTGAAATGTCCTGCTGATATGTTAAAGAGGCAAGT 2580
Qy      1849 GTGAAGATCCTAGCTGCTGCTGCTGCAATTTGCAATTTTCACTGTCAGCAGAGAT 1908
Db      2581 GTGAAGATCCTAGCTGCTGCTGCTGCAATTTGCAATTTTCACTGTCAGCAGAGAT 2640
Qy      1909 TTTCTTAATAAGAGCTCATCTTTTCAAGCAGAGATATGTTCTTCTATGAGATCTGAT 1968
Db      2641 TTTCTTAATAAGAGCTCATCTTTTCAAGCAGAGATATGTTCTTCTATGAGATCTGAT 2700
Qy      1969 GTGCTTACCA----- 1978
Db      2701 GTGCTTACCAATGATGATTAATATGATGCAATTTTCATATATCTGCAATGCTCAAAATA 2760
Qy      1979 ----- 1978
Db      2761 TGCTTTGTTGTGAGCTTAAGAACATATGTTCCACTTAATACATGTGCCAAAAGTTGTAAC 2820
Qy      1979 ----- 1978
Db      2821 AAGATTAAACAAGTGTCTGATTAATTTTCACTAATATATGCTGCTGAATTTTTCATCAA 2880
Qy      1979 ----- 1978

```

```

Db      2881 CTGTAGACAGAAATGTAAATTTCACTCTCAACATTTCTGTTAGATAAGTAGATTAG 2940
Qy      1979 -----T 1979
Db      2941 AGATTGCTTAGTGTGCTTGTGCTTGTCCAACTTTTCTTCTGATTTTCTTTTCGATTT 3000
Qy      1980 AGGTCAGTCAAGCTGACGATTCAGAGAGCACTTCCAGAAATGATGCTAGAGACTGACAGA 2039
Db      3001 AGGTCAGTCAAGCTGACGATTCAGAGAGCACTTCCAGAAATGATGCTAGAGACTGACAGA 3060
Qy      2040 GAATTAATATCCAAAGTGGCAGAAATTAATGCTCTGCTTTGGGCTGATACCCGAT 2099
Db      3061 GAATTAATATCCAAAGTGGCAGAAATTAATGCTCTGCTTTGGGCTGATACCCGAT 3120
Qy      2100 AGAAATGTTACCAAG----- 2113
Db      3121 AGAAATGTTACCAAGAGTGGAGGAATTAATCTACAAATTCATCAATTTGTGAAAACCTGT 3180
Qy      2114 -----AGTTTTGATGGGC 2128
Db      3181 TGACATGATTAATAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3240
Qy      2129 GAATGCTGAAGATTTTGAATGACAGAGAGCTGAAACCTGCCAGCTTGGTTGTTATG 2188
Db      3241 GAATGCTGAAGATTTTGAATGACAGAGAGCTGAAACCTGCCAGCTTGGTTGTTATG 3300
Qy      2189 ATTAATACCTGTTGAACCTATCTGTTGACAGTGTGACAGCTGACAGATGGAACCCGTG 2248
Db      3301 ATTAATACCTGTTGAACCTATCTGTTGACAGTGTGACAGCTGACAGATGGAACCCGTG 3360
Qy      2249 CTCTGTGGAAGCACTCTGAGAGAGTCTGCTTCTATCTGATTTGTTGTTTATCCAGAA 2308
Db      3361 CTCTGTGGAAGCACTCTGAGAGAGTCTGCTTCTATCTGATTTGTTGTTTATCCAGAA 3420
Qy      2309 ACAATGCTACTGATGTGAGAACTTACACACACAAATAGAGAAATTTCTGTGTCAAGTACG 2368
Db      3421 ACAATGCTACTGATGTGAGAACTTACACACACAAATAGAGAAATTTCTGTGTCAAGTACG 3480
Qy      2369 GGTGAAAATCACTGAAGGCTCTGTTCTTGATCATATA 2406
Db      3481 GGTGAAAATCACTGAAGGCTCTGTTCTTGATCATATA 3518

RESULT 7
US-10-600-070-10
; Sequence 10, Application US/10600070
; Publication No. US20040139500A1
; GENERAL INFORMATION:
; APPLICANT: Osteoryoung, Katherine W.
; APPLICANT: Viltha, Stanislaw
; APPLICANT: Kokharova, Olga A.
; APPLICANT: Gao, Honoo
; TITLE OF INVENTION: Placid Division and Related Genes and Proteins, and Methods of
; FILE REFERENCE: Use
; CURRENT FILING DATE: 2003-06-20
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 10
; LENGTH: 3667
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-600-070-10

Query Match 71.5%; Score 1720.8; DB 7; Length 3667;
Best Local Similarity 79.1%; Pred. No. 0;
Matches 2404; Conservative 0; Mismatches 2; Indels 632; Gaps 5;

Qy      1 ATGAAGCTCTGAGTCAAGTCCGATGCTCTCCCAATTCATTAATGATCCGATTAACA 60
Db      481 ATGAAGCTCTGAGTCAAGTCCGATGCTCTCCCAATTCATTAATGATCCGATTAACA 540

```

QY	61	CCGGGAGAGCAAAAGCTCCGAGCTGACCAACAACACTTACATCATCTGTCTCCGACAC	120
Db	541	CCGGGAGAGCAAAAGCTCCGAGCTGACCAACAACACTTACATCATCTGTCTCCGACAC	600
QY	121	AAATGGGCGACCGCTCTCTCGAGCTTCAATTCACCTCGATTCCTCCCTCCCTCC	180
Db	601	AAATGGGCGACCGCTCTCTCGAGCTTCAATTCACCTCGATTCCTCCCTCCCTCC	660
QY	181	TTTCGCACCGCCACCGACCGGCACTCTCGTCTCTGCGACCACTCATTTGATGCTGCC	240
Db	661	TTTCGCACCGCCACCGACCGGCACTCTCGTCTCTGCGACCACTCATTTGATGCTGCC	720
QY	241	GAAGCCGACGTCCTCCATCCCATTTGATTTCTACACAGGTAATTAGAGCTCAAAACATTC	300
Db	721	GAAGCCGACGTCCTCCATCCCATTTGATTTCTACACAGGTAATTAGAGCTCAAAACATTC	780
QY	301	TTAACCGATGGATCAGAAAGACATTCGAGAGCTAGGTTTGAAACCGCGGCATTCGGGT	360
Db	781	TTAACCGATGGATCAGAAAGACATTCGAGAGCTAGGTTTGAAACCGCGGCATTCGGGT	840
QY	361	TTTCAGCGACGACGCTTTTAATCAGCCGAGACAGATTCCTTCAAGCTGCTTGGAAACTCTG	420
Db	841	TTTCAGCGACGACGCTTTTAATCAGCCGAGACAGATTCCTTCAAGCTGCTTGGAAACTCTG	900
QY	421	TCTATATCTCGGTCTAGAAAGAGATCAATGAAAGTCTTCTTGATGATGAAGAGCTACA	480
Db	901	TCTATATCTCGGTCTAGAAAGAGATCAATGAAAGTCTTCTTGATGATGAAGAGCTACA	960
QY	481	GTCAATCATGATGTTCTCTGGGAT-----	504
Db	961	GTCAATCATGATGTTCTCTGGGATTAAGTAATTCGATTCGGAATATTAAGTTTCTTC	1020
QY	505	-----AAGTTCTCTGGGAC	518
Db	1021	GTTTTAATTCATGAATTGATTAAGAAAGAACTTTATCTAGTAAGAGTTTCTCGGGGC	1080
QY	519	TCTCTGTGTATTTGCAAGAGGTGGTGAGACTGAGATAGTTCTTGGGGTGGTGAGGCTCT	578
Db	1081	TCTCTGTGTATTTGCAAGAGGTGGTGAGACTGAGATAGTTCTTGGGGTGGTGAGGCTCT	1140
QY	579	GCTTAAGAGAGGTGGCTAGTGGTTTAAGCAAGATGGGTTTGTAGTTAGTGAAGGGCTTGC	638
Db	1141	GCTTAAGAGAGGTGGCTAGTGGTTTAAGCAAGATGGGTTTGTAGTTAGTGAAGGGCTTGC	1200
QY	639	GTTCCTCGATGCTTCGAGGAGTGCATAGGACATTTGGATCACTGATTTTATTAATCTGTGA	698
Db	1201	GTTCCTCGATGCTTCGAGGAGTGCATAGGACATTTGGATCACTGATTTTATTAATCTGTGA	1260
QY	699	TGAGTTTGTAGAGAGCTTGAAGCTTTTAC-----	730
Db	1261	TGAGTTTGTAGAGAGCTTGAAGCTTTTACAGGTAAGTTTGAAGCTTGTGTAAATTTG	1320
QY	731	-----A	731
Db	1321	ACGAGCGTTGCTTTAATAAGAACTTTCCTGATTGATACCTTGTGATTAAGTCTGTGTGTA	1380
QY	732	GGAGGAAAGAGCAAGTACCTCTGCAACCGGATTTTACGTGACAAATTTGATGAGCTTTGGA	791
Db	1381	GGAGGAAAGAGCAAGTACCTCTGCAACCGGATTTTACGTGACAAATTTGATGAGCTTTGGA	1440
QY	792	AGAGATCACTCCGCGGTATAGTCTTGGAGCTACCTTGGCTTACCGCTTGGTATGATTAAGC	851
Db	1441	AGAGATCACTCCGCGGTATAGTCTTGGAGCTACCTTGGCTTACCGCTTGGTATGATTAAGC	1500
QY	852	TGCGAAAAGACTTAAATGGTTTAAAGCGGTGTGCGAATTTTGTGTGCTTGTGAGAGAG	911
Db	1501	TGCGAAAAGACTTAAATGGTTTAAAGCGGTGTGCGAATTTTGTGTGCTTGTGAGAGAG	1560
QY	912	TGAGAGCATCAGCTCTTGTGTGGGGGTTTAAACCGGTAGAAAGTTTATGATGAGGGGCTTTT	971
Db	1561	TGAGAGCATCAGCTCTTGTGTGGGGGTTTAAACCGGTAGAAAGTTTATGATGAGGGGCTTTT	1620
QY	972	ACGAATGACAGCTGCTGAGC-----	991

Db	1621	ATGATAGACAGCTCTGAGCAGGTAAACAGTTAGATACCTTTTAAATTTCTTANCA	1680
Qy	992	-----AGTTGATCTTTTGTGA	1008
Db	1681	TGATATAACTTTAGGTTTCATTTTAAATGATGTGTGTGATAGGTTGATCTTTTGTGA	1740
Qy	1009	GCTACCCCAAGCAATATTCACGACAGCTCATTTGAAGTTTACGAAGTTGACCTTCTCT	1068
Db	1741	GCTACCCCAAGCAATATTCACGACAGCTCATTTGAAAGTTTACGAAGTTGACCTTCTCT	1800
Qy	1069	GTGGCTCAAGCTTTATTTGTTAATAACCAACCTTTTACAGATGCTGATPAGCAATTC	1128
Db	1801	GTGGCTCAAGCTTTTATTTGTTAATAACCAACCTTTTACAGATGCTGATPAGCAATTC	1860
Qy	1129	CAGCAACTTCACGACAGCTAAGGTATAGGCTATGAGATTCCTGCGATGTTGATATGATACA	1188
Db	1861	CAGCAACTTCACGACAGCTAAGGTATATGGCTATGAGATTCCTGCGATGTTGATATGATACA	1920
Qy	1189	CGGAATATATGGAGATAGACCTTCGGTCTAGAAAAGGGAACCTGTGACCTGTTATPAGGC	1248
Db	1921	CGGAATATATGGAGATAGACCTTCGGTCTAGAAAAGGGAACCTGTGACCTGTTATPAGGC	1380
Qy	1249	AAAGTTGATGAATCCCGTATGTGGTTGGGCTTACACGTGAGATTCACAAATPAGGAAT	1308
Db	1981	AAAGTTGATGAATCCCGTATGTGGTTGGGCTTACACGTGAGATTCACAAATPAGGAAT	2040
Qy	1309	CCAGCTATTTGAGAGTTTGTTTTGGAAATTCAAATTCGTATGACAAATGATGATCTCCCT	1368
Db	2041	CCAGCTATTTGAGAGTTTGTTTTGGAAATTCAAATTCGTATGACAAATGATGATCTCCCT	2100
Qy	1369	GGACTATGCAAAATGTTGGAAAACCTGTGTGGCAGAGGGTGTCTTCTCTAGGTTACAGAC	1428
Db	2101	GGACTATGCAAAATGTTGGAAAACCTGTGTGGCAGAGGGTGTCTTCTCTAGGTTACAGAC	2160
Qy	1429	ACCAAAAGTTAAAAATTTAACTCGGGGACATACATGATATGATCTTATGTTTGAAGTTAC	1488
Db	2161	ACCAAAAGTTAAAAATTTAACTCGGGGACATACATGATATGATCTTATGTTTGAAGTTAC	2220
Qy	1489	TTGGAAAAGTGGAGGTAGTTCAAGGGTCTTCCTTTAGCTGCTGCTGACATATGAGCAAGG	1548
Db	2221	TTGGAAAAGTGGAGGTAGTTCAAGGGTCTTCCTTTAGCTGCTGCTGACATATGAGCAAGG	2280
Qy	1549	ATTGAGCCGACGACGTGTGAAGCTTAGTCTATGACAGGCACTGCAGAAAGTTTCTCTCC	1608
Db	2281	ATTGAGCCGACGACGTGTGAAGCTTAGTCTATGACAGGCACTGCAGAAAGTTTCTCTCC	2340
Qy	1609	CGCTATACAGATGAAAACCTGGGCTGAAACCCAAGATGTGCAGAGACAGTCTTAACTGTA	1668
Db	2341	CGCTATACAGATGAAAACCTGGGCTGAAACCCAAGATGTGCAGAGACAGTCTTAACTGTA	2400
Qy	1669	GATCCGTGGGTAAACAATGTAGGCGGATAGGTGAGGCTGAGTGTCTTATTTGCAAGAACT	1728
Db	2401	GATCCGTGGGTAAACAATGTAGGCGGATAGGTGAGGCTGAGTGTCTTATTTGCAAGAACT	2460
Qy	1729	GTAAGACCCCTCTGAAAACTTTGAAAACCTAATGATTATGCAATTCGAGCTGGGGTCTCGAG	1788
Db	2461	GTAAGACCCCTCTGAAAACTTTGAAAACCTAATGATTATGCAATTCGAGCTGGGGTCTCGAG	2520
Qy	1789	AGTAGCGTTGATGAACCTACTGTTGAAATGTCCGTTGCTGATATGTTAAAGAGGCAAGT	1848
Db	2521	AGTAGCGTTGATGAACCTACTGTTGAAATGTCCGTTGCTGATATGTTAAAGAGGCAAGT	2580
Qy	1849	GTGAAGATCCAGTGTGCTGGGTGGGCAATTGACATGTTTCACTGTTCAACCCAGAAAGTAT	1908
Db	2581	GTGAAGATCCAGTGTGCTGGGTGGGCAATTGACATGTTTCACTGTTCAACCCAGAAAGTAT	2640
Qy	1909	TTTCTTAAAAAGCACTCATCTTTTCAACGCAAGGATATGTTTCTTCTATAGGAATCTGAT	1968
Db	2641	TTTCTTAAAAAGCACTCATCTTTTCAACGCAAGGATATGTTTCTTCTATAGGAATCTGAT	2700
Qy	1969	GTGCGTACCA-----	1978

```
Db 2701 GTGCTACCATAGATGATTAATGATGCAATTTTCATATATCTGACTTCTCAAAAT 2760
Qy 1979 ----- 1978
Db 2761 TCGTTGTTTGTAGAGTAGAACAATAGTCCACCTTAATACATGTCCAAAGTTGATCC 2820
Qy 1979 ----- 1978
Db 2821 AAGATTACAAAGTGTCTGAGTAATTTCACTAATATGCTGCTGAATTTTGTATCAA 2880
Qy 1979 ----- 1978
Db 2881 CTGTAGACAGAAATGTAATTTCACTCTCAACATTTCTGTTAGATTAACGTAGATTAG 2940
Qy 1979 ----- 1979
Db 2941 AGATTGCTTAGTGTGCTTTGTCCAACTTTCTTTCCATGATTTTCTTTTGATTT 3000
Qy 1980 AGGTCAGTCAGAGCTGACGATTCAGAGCACTTCCAGAAATGATGCTAGAGCTGACAG 2039
Db 3001 AGGGTCAGTCAGAGCTGACGATTCAGAGCACTTCCAGAAATGATGCTAGAGCTGACAG 3060
Qy 2040 GAATATAGATCAAGTCGACAGAGTTAAGTCTGCTGCTTTGGGCTTGATCAGCCGAT 2099
Db 3061 GAATATAGATCAAGTCGACAGAGTTAAGTCTGCTGCTTTGGGCTTGATCAGCCGAT 3120
Qy 2100 AGAAATGTTACCG----- 2113
Db 3121 AGAAATGTTACCGAGGTGAGGGAATTAATCTACAATTCATGATGTCAGAAACTGT 3180
Qy 2114 -----AGGTTTGAATGAGC 2128
Db 3181 TGGACATGATTAATAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3240
Qy 2129 GAATGCTGAAGTTTGAAGCTGACAGAGCACTGAAAGCTGAGGCTGGGCTGCTGCTGCT 2188
Db 3241 GAATGCTGAAGTTTGAAGCTGACAGAGCACTGAAAGCTGAGGCTGGGCTGCTGCTGCT 3300
Qy 2189 ATTATACACTGTTGAACATCTATCTGTGACAGTGTGACAGTCTCAGACAGATGGAACCCGTG 2248
Db 3301 ATTATACACTGTTGAACATCTATCTGTGACAGTGTGACAGTCTCAGACAGATGGAACCCGTG 3360
Qy 2249 CTCTGTGGAAGCACTGTGAGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2308
Db 3361 CTCTGTGGAAGCACTGTGAGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3420
Qy 2309 ACAATGCTACTGATGTCAGAACCTTACACAACAAGATAGAGGTTTCTGCTGCTGCTGCTG 2368
Db 3421 ACAATGCTACTGATGTCAGAACCTTACACAACAAGATAGAGGTTTCTGCTGCTGCTGCTG 3480
Qy 2369 GGTGGAATAATCACTGAGGCTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2406
Db 3481 GGTGGAATAATCACTGAGGCTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3518

RESULT 8
US-10-600-070-132/c
; Sequence 132, Application US/10600070
; Publication No. US20040139500A1
; GENERAL INFORMATION:
; APPLICANT: Osteeryoung, Katherine W.
; APPLICANT: Vitsha, Stanislaw
; APPLICANT: Kokeharova, Olga A.
; APPLICANT: Gao, Hong
; TITLE OF INVENTION: Placid Division and Related Genes and Proteins, and Methods of
; FILE REFERENCE: MSU-08153
; CURRENT APPLICATION NUMBER: US/10/600,070
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 132
; LENGTH: 561
```

```
TYPE: DNA
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: misc feature
LOCATION: (127)..(127)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (520)..(520)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc feature
LOCATION: (541)..(541)
OTHER INFORMATION: n is a, c, g, or t
US-10-600-070-132

Query Match 20 0%; Score 482; DB 7; Length 561;
Best Local Similarity 99.4%; Pred. No. 1e-134;
Matches 482; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1922 GTCATCTTTTCAACGCAAGATAGGTTTCTCTATGGAATCTGATGCTACCATAG 1981
Db 561 GCTCATCTTTTCAACGCAAGATAGGTTTCTCTATGGAATCTGATGCTACCATAG 502
Qy 1982 GGTCAATCAGAGCTGACGATTCAGAGCACTTCCAGAAATGATGCTAGAGCTGACAG 2041
Db 501 GGTCAATCAGAGCTGACGATTCAGAGCACTTCCAGAAATGATGCTAGAGCTGACAG 442
Qy 2042 ATATATGATCCAAATGAGGAGAAATTAAGTCTGCTGCTTTGGGCTTGATCAGCCGAT 2101
Db 441 ATATATGATCCAAATGAGGAGAAATTAAGTCTGCTGCTTTGGGCTTGATCAGCCGAT 382
Qy 2102 AAATGTTACAGAGGTTTGGATGGGCAATGCTGAAGATTTGACAGAGCAAGCTG 2161
Db 381 AAATGTTACAGAGGTTTGGATGGGCAATGCTGAAGATTTGACAGAGCAAGCTG 322
Qy 2162 AAATGCGCAGCTTGGGCTGTTATATATTAATTAACACTGTTGAAACTATCTGTTGACAGTG 2221
Db 321 AAATGCGCAGCTTGGGCTGTTATATATTAATTAACACTGTTGAAACTATCTGTTGACAGTG 262
Qy 2222 TGAACGCTCAGCAGATGGAACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2281
Db 261 TGAACGCTCAGCAGATGGAACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 202
Qy 2282 GTCTATCTGATTTGCTGCTATCAGAAAAAATGCTACTGATGTCAGAACTTACACAAC 2341
Db 201 GTCTATCTGATTTGCTGCTATCAGAAAAAATGCTACTGATGTCAGAACTTACACAAC 142
Qy 2342 GATACGAAGTTTCTGCTCAGAGTCAGGCTGGAATAATCACTGAAGGCTCTGCTTGCAT 2401
Db 141 GATACGAAGTTTCTGCTCAGAGTCAGGCTGGAATAATCACTGAAGGCTCTGCTTGCAT 82
Qy 2402 CATTA 2406
Db 81 CATTA 77

RESULT 9
US-10-600-070-126
; Sequence 126, Application US/10600070
; Publication No. US20040139500A1
; GENERAL INFORMATION:
; APPLICANT: Osteeryoung, Katherine W.
; APPLICANT: Vitsha, Stanislaw
; APPLICANT: Kokeharova, Olga A.
; APPLICANT: Gao, Hong
; TITLE OF INVENTION: Placid Division and Related Genes and Proteins, and Methods of
; FILE REFERENCE: MSU-08153
; CURRENT APPLICATION NUMBER: US/10/600,070
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: PatentIn version 3.2
```

SEQ ID NO 126
LENGTH: 2283
TYPE: DNA
ORGANISM: Oryza sativa
US-10-600-070-126

Query Match 20.0%; Score 481; DB 7; Length 2283;
Best Local Similarity 54.3%; Pred. No. 5,2e-134;
Matches 1256; Conservative 0; Mismatches 880; Indels 177; Gaps 7;

108 CTGCTCCGCGCAGAAATGGCGCCGACCGCTCTCTCCGACCTTCAATTCACCTCCGATTC 167
132 CTGGGCGCCAAAGCCTCTTCCGCGACTTCCACCTCTCCGACCGCGCCCTCCGACCC 191
168 CTCCTCTCTCTCTCTGCGCACCGCACCACCGCACACTCTGCTCTCTGCGACATC 227
192 GCCGTCCCGCGCCCGCGCCCGCGCCCGCCCTCCGCTCCGCTCCGCTCTGCTCTT 251
228 TATTGATCGTCCCGACCGCAGCTCCGATCCCATTTGATTTCAACAGTATTAGAGC 287
252 CCGCGACGCGCGGACGCTCCCTCCGCTCCAGTCGATTTCTACAGGTTCTAGGGGC 311
288 TCAACACATTTCTTACCGATGATGAAAGAGCATTCGAGCTAGGGTTGAAAC 347
312 AGAGCCACATTTCTTGGCGATGCGATCAGAGGCGGCTTCGAGCGCAGATAGCCAC 371
348 GCGCGATTCGAGTTTCAAGCAGCAGCTTTATCAGCGGAGACAGATTTCTTCAAGCTC 407
372 ACCGAGATAGCTACAGCAGGATGCTCTGTTGTTGCGACAAATGCTGCGATTC 431
408 TTGCGAACTCTGCTTAACTCTCGGCTTGAAGAGATGAAATGAGGCTCTTCTGATGA 467
432 CCATGACACTCTCATGAACAGAACTCCGCACTCAGATGATCGTGGCTTTCTAGAA 491
468 TGAAGAACTTACGATCACTGATGTTCTTGGGATGATGTTCTTGGGGCTCTCTGCT 527
492 CCGTGAAGAGCTCTCACATGATATGTTGTTGGAAGAGCGCTGGG----- 540
528 ATTGCAAGAGGTGGTGAAGTGAATGTTCTTGGGTTGGTGAAGCTCTGCTTAAAGA 587
541 -----GAGGCACTTGTGCTTGTAACTGAGAGAAAGCTTCTTCTGGA 584
588 GAGTTGCTTAAGTGTAAAGCAAGATGAGTTTAACTTAACTGCGCTTGGCTTCTGA 647
585 TCGGCACTTCAAGCGCTTCAAGCAGATGCTGCTGATGATGCTGCTTCTGATGGA 644
648 TGTCTCGAGGATGCTATGCGATTCACCTGATTTTATTTACTGTTATGAGTTGT 707
645 TCTATCAAGGATGCTATGCGAGCAAGCCCTCCAGATGATTAATGCTCTGCGAGTGT 704
708 TGAGGAAGCTTTGAAGCTTTTACAGAGGAAGAGCAAGTACCTTGCACCGGATTTAG 767
705 CGAGAGGCTCTCAAGCTCTTGCAGAGAGATGAGCAAGCAATCTCGACCTGATCTCT 764
768 TGCAAAATTTGATGAGACTTTGGAAGAGATCACTCCGCTTATGCTTGAAGCTTGG 827
765 TTCAAGATTTATGAATCTCTGAGAGATTTACACTCGCTGTTATTTGAGCTCTCTC 824
828 CTTACCGCTTGTGATGATGATGCTGCGAAAAGATTAATGTTTAAAGCGTGTGCGAA 887
825 CTTTCTTATTTGACACAGAGATCATTAAGAGCGCCAAAGGCTTCAAGGTGCGAAGAA 884
888 TATTTTGATCTGTTGAGAGAGTGAAGCATCAGCTCTTGTGGGGGTTTGAACCGGGA 947
885 CATTTTGTGAGCGTTGGCAGAGAGATTTGCTACCGTTGAGAGAGATTTTCTCGTGA 944
948 GAAGTTTATGATGAGGCGTTTATTCAGATGACAGCTGCTGAGCAGGTTGATCTTTTGT 1007
945 AGCTTTATGATGAGGCTTTTATTTGAGATGACATCAATGAAAGATGATTTCTTTTC 1004
1008 AGCTTACCCCAAGCATATTTCCAGCAGATCATTTTGAAGTTTACGAAGTTGCACTTGTCT 1067
1005 AAAAAACCGAATAGCATTTCTCTGAAATGTTGAAATTTTACATTTATGACATTTGACA 1064

1068 TGTGCTCAAGCTTTTATTTGTAAGAAGCACACTTTTACAGAGATGATTAAGCAAT 1127
1065 TGTGCTCAAGCATTTTATGTAAGGCGCACATTTATCATATGCGGATGATCTTTT 1124
1128 CCAGCAACTTACAGGCTTAAGGTAATGCTATGAGATTCCTGCGATGTTGATGATAC 1187
1125 TGAACCACTCCAGAAAGTTCAA-----CATAGGTTCTCATTA 1160
1188 ACGGAATATTGGAGATTAACCTTCCGCTTGAAGGAGGACCTGTCACCTGCTTAAG 1247
1161 TGCTTATGATTAATGAGATGAGCTTGCATTTGAAAGGCGATTTCTGCTCATTTGCTGG 1220
1248 CAAAGTTATGAAATGCGGATGTTGTTGGCTTGAACAGTGAAGATTCACAAATATGAA 1307
1221 AGATTTAGCAAGTGCAGAAATGTGCTTGAATTTGATTAATGAGCTTCAACATACAGAA 1280
1308 TCCAGCTATTGTGAGGTTGTTTGAAGAAATTCAAATC--GTATGACATATGATCT 1364
1281 CCGCAAAATTTCTAGAGTTTATTTGACCAACTAGCATCAGTGAAGAGATGATCTTCT 1340
1365 CCGTGAATATGCAAAATTTGTAAGAACTGTTGGACGAGGGTGTCTTCTTGAAGTTAG 1424
1341 TCCAGGCTGTGCAAGCTTTTGAAGACTTGGCTTATCTTGAAGTTTCTTGAAGAGAG 1400
1425 AGACACCAAGATTTAAATTTTAACTCGGGGACTACTATGATGATCTTATGCTTTTGAAG 1484
1401 AGATACCTCGGGGATGAGATGTTCAAGCTTGAAGATTAACATGATGATCAAGATTTTAA 1460
1485 TTACTTGAAGAGTGAAGTGAATGTTCAAGGTTCTCTTAACTGCTGCTGCAATATGCT 1544
1461 CTACCTTGAAGAGATGAGGAGTGTGCTCTTCAATTTGGCTGCTGCTCTCTATATTC 1520
1545 AAGGATTTGAGCGGAGATGAAAGCTAGTGCATGACAGGACCTGCAAGAAAGTTTTC 1604
1521 AAAAATTGTGCTCAAGCTACAGCTGCACTTGGTA----- 1555
1605 TTCCCGCTATACAGATGAAACTCGGCTGAACCAAGATGTGCAAGAGACAGTGTTAAG 1664
1556 -----CTGTGAATTAATATGATTAAGCG----- 1581
1665 TGTATGCTGTTGTTGTAACAATGATGAGCGGTGATGTGAGCTGTGTCTTTATTTGAGA 1724
1582 -----TTCAACAA 1589
1725 AGCTGTAAGCCCTGTAAGAACTTTGAACCTTAATGATTAATGCAATTCAGCTGGGCTCT 1784
1590 GGTTTTCCATTGATTAACAAGTTAGCAGGTGACCATGAAATATCTAAAGATGGCCC 1649
1785 AGAGATGAGGTTGATGAACTACTGTTGAAATGTCGTTGCTGATATGTTAAAGAGGC 1844
1650 TGGGGAATCTTGAATAATTTTGAACAGGAAATGACCTGCTCATATGATTCGAAATGC 1709
1845 AAGTGAAGATCTTACGCTGCTGTGTGCGAATTTGACATGATTTCACTGTTCAAGCAGAA 1904
1710 CGCCTTGAATTAATCTCTGCTGCGCATGTTTGAACCTGTTGCAAGTAAATGAGGCGCAA 1769
1905 GATTTTCTTAAAGCAGCTCATCTTTTCAACCGAAGATATGTTCTTATGATC 1964
1770 ATATTTGCTCTGTAAGGCGCCCTTCTGCTATTTGAGATGAGCATGATCTGTGAGCT 1829
1965 TGAT-----GTGCTACATAGGCTGATGAGTCAAGCTGACGATTCAGAAACCTTCC 2015
1830 TGCTAATAGTGTGACCTTACTGATGATCTGCACTAGATGAAGATCAGATACATATTC 1889
2016 CAGAAATGATGCTAGACTGCAAGAAATATATGATTCGAATGCGCAAGATTAAGTCTT 2075
1890 TAGAATGATGCGAAGCTGCGAAGATATTTGTCGCAAGTGCAGAGTATCAATTA 1949
2076 GGCCTTTGGGCTGATCAGCGCATAGAAATGTTACGAGAGTTTGGATGGGCGAATGCT 2135
1950 GGCCTTGGAGCAGAACATTTGCTGATCATTTGCAAGAGGTTCTTATGCAACATGCT 2009

QY	2136	GAAGATTTGGACTGACAGAGAGCTGAAATCGCGAGCTTGAGTTGGTTTATGATTATAC	2195
QY <td>2010</td> <td>AAAGGTGTGACCTACCGACGACGGAGATTGAGCGTATGGGTGTTCTTGGAGATATAC</td> <td>2069</td>	2010	AAAGGTGTGACCTACCGACGACGGAGATTGAGCGTATGGGTGTTCTTGGAGATATAC	2069
QY <td>2196</td> <td>ACTGTGAACTATCTGTGTGACAGTGTGACAGTCTCAGCAGATGAGAACCGGTGCTTGGT</td> <td>2255</td>	2196	ACTGTGAACTATCTGTGTGACAGTGTGACAGTCTCAGCAGATGAGAACCGGTGCTTGGT	2255
Db <td>2070</td> <td>ACTATCGGATGTGACGATTGATGATCATCATCTCCCTTAAGATGGTTCGACGACGACATGT</td> <td>2129</td>	2070	ACTATCGGATGTGACGATTGATGATCATCATCTCCCTTAAGATGGTTCGACGACGACATGT	2129
QY <td>2256</td> <td>GGAAGCAACTGTGAGGAGTCTGGCTGTGTCATCTGATTGTGTTATCCAGAAAACAATGC</td> <td>2315</td>	2256	GGAAGCAACTGTGAGGAGTCTGGCTGTGTCATCTGATTGTGTTATCCAGAAAACAATGC	2315
Db <td>2130</td> <td>GGAGGCTTACGATTATATGAGCAGGCGCAACTTACTGATTTTACTGAGCCCAAGAAACAATGA</td> <td>2189</td>	2130	GGAGGCTTACGATTATATGAGCAGGCGCAACTTACTGATTTTACTGAGCCCAAGAAACAATGA	2189
QY <td>2316</td> <td>TACTGATGTGGAACCTTACACAAACAAGTATTCGAAAGTTTCTGTTCCAA---TCAGGGTGG</td> <td>2372</td>	2316	TACTGATGTGGAACCTTACACAAACAAGTATTCGAAAGTTTCTGTTCCAA---TCAGGGTGG	2372
Db <td>2190</td> <td>TTCTATATGACACAATAATCACTACCCGGTATGAGATGGCTTCTTCCAAAGTAAAGGGTG</td> <td>2249</td>	2190	TTCTATATGACACAATAATCACTACCCGGTATGAGATGGCTTCTTCCAAAGTAAAGGGTG	2249
QY <td>2373</td> <td>GAAGATCACTGAGAGGCTCTGTTCTTTCGATCTACTA</td> <td>2405</td>	2373	GAAGATCACTGAGAGGCTCTGTTCTTTCGATCTACTA	2405
Db <td>2250</td> <td>GAGATTAACGAGAGAGACAGTCTTCCAAAGTCGTA</td> <td>2282</td>	2250	GAGATTAACGAGAGAGACAGTCTTCCAAAGTCGTA	2282

RESULT 10
US-10-424-599-129007

Sequence 129007, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 129007
LENGTH: 1146
TYPE: DNA
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: clone ID: PAT_MRT3847_87499C.1
US-10-424-599-129007

Query Match	19.7%	Score 472.8	DB 7	Length 1146
Best Local Similarity	70.5%	Pred. No. 1e-131		
Matches 649, Conservative	0	Mismatches 262	Indels 9	Gaps 1

QY	107	TCGACTCCGCAGCAAAATGGGCGCAGCCGTCCTCTCCGCACTTCAAAATTCACCCCGAAT	166
Db	236	TATCCGCACACAGCAAAATGGGCGAGGCGCTCATTCGCGATTTCCAAATTCCTCGGCGACG	295
QY	167	CCCTCCTCCTCCTCTTCGCGCACCGGCACACACACCGCGACCTCTCGTCTCTTCGCACCAT	226
Db	296	CCGCGCGCTCOACTT-----CGACCTCGACTCTCAGCCCTCTCCGCTCCTCC	346
QY	227	CTAATTGATCGTCCCGAAAGCGACGTCCTCCCATTCGATTTCTACCAAGTATTAGAG	286
Db	347	GCCCTCATCTCCGACGCGCTACGTCGTGATCCCTCTCGACCTGTACCGCATTCCTCGGCG	406
QY	287	CTCAAAACAATTTCTTAAACGATGGATTCAGAAAGATTCGAAGCTAGGGTTTCGAAC	346
Db	407	CCGAACGCACTTCTCTGGGCGACGCAATTCGCAAGGCTTCAGAGGCCAAAGTTCTCGAAGC	466
QY	347	CGCGCAATTCGGTTTCAGCGACGACGCTTTAATCACGCCGAGACAGATTCCTCAAGCTG	406
Db	467	CTCTCTAGTAGGCTTCAGCAACGACGCTTTAATCACGCCGCCCAATCTCTCAACAG	526
QY	407	CTTCGCAAACTCTGTCTAATCTCTGATCTAGAAAGAGTCAATGAAGGTCCTCTTGATG	466
Db	527	CTGCGAAACCTTAGCTGATCTCTACTTCGAAAGAGGTTCAAATCAAAAGCTCTGTGACG	586
QY	467	ATGAGAGACTACATCATCTGATGTTCTCTGGATTAAGTTCCTGCGGCTCTCTGTG	526

Db	587	ACGAGAAAGCCGCATTCTCACTAAATCCCTTTTCGACAAAGTCTCGAGACCGTTGTGTC	646
Qy	527	TATTCGAAGAAGTGTGTAGACCTGAGATAGTATCTTCGGTGTGTGAGGCTCTGCTTAAG	586
Db	647	TGTTGCAGGAAGCTCGAGAGACGAGAGCTGTGTGCTTGAAGATTGGCAGGGTTTGCTTAGGG	706
Qy	587	AGAGGTTCCTTAAGTCCGTTTAAAGCAAGATGTGTTTAGTTAGTGGCCCTGCTTCG	646
Db	707	AGAGGTTCGCGAAGACGTTTAAAGCAGATGTGTGTGCTATAGGCACTCGCATTTGTG	766
Qy	647	ATGTCTGAGGAGATGCTATGGCATTTGATCCACTGATTTTATTACTGTGTATGAGTTTG	706
Db	767	ACGGTCAAGGAATGCTATGAGCCCTTGCTCTCACCGGATTTTCATTTGCGCCTGTGAGATGC	826
Qy	707	TTGAGGAAGCCTTGAACCTTTTACAGAGGAAGGAGCAAGTAAAGCCCTTGACACCGGATTAAC	766
Db	827	TCGAGAGGGGCTTGAACCTTTTGTGAGGAGAAAGGGGAGAACAGCTAGCTCAGATTAAC	886
Qy	767	GTGACACAAATTTGATGACATTTTGGAGAGATCACTCCGCTTATATGCTTGAAGCTTAG	826
Db	887	AAGCACAATATGATGAGACCGCTAGAGAGATTAACCCACGTTGTGTTTGGAACTTTAG	946
Qy	827	GCTTACCGCTTGTGATGATTAACGCTGCGAAAAGACTTAAATGTTAAGCGGTGTGCGGA	886
Db	947	CTTGTCCTCTTGATGACGAACATGACAGCGGAGGAGGAGAAAGTCTTCTTGTTGTCCGTA	1006
Qy	887	ATATTTTGTGTGCTGTGTGAGAGAGGTGAGACATCAGCTTGTGTGGGGGTTTGACCCGTG	946
Db	1007	ACATTTTGTGTGGGCGGTTGTGTGAGGGGGGTGCACCAATATGCGCGGGGTTTACCCGTG	1066
Qy	947	AGAAGTTATGATGAGAGCGCTTTTACGAATGACAGTGTGAGCAGGTGATCTTTTTG	1006
Db	1067	AAGCTTTCATGAATGAGAGCATTTCTTACATGACAGCGCTTGAACAGTTGAACCTTTTTG	1126
Qy	1007	TAGCTACCCCAAGCAATATT	1026
Db	1127	TAGCCACCAAGTACTATT	1146

RESULT 11
US-10-425-115-81853

```

Sequence 81853, Application US/10425115
Publication No. US20040214272A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovacic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants
FILE REFERENCE: 38-21 (53222) B
CURRENT APPLICATION NUMBER: US/10/425.115
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 369326
SEQ ID NO 81853
LENGTH: 1411
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: clone ID: MFT4577_174665C.1
US-10-425-115-81853

```

Query Match	15.0%	Score	361.4	DB	8	Length	1411
Best Local Similarity	61.3%	Pred. No.	7.4e-98				
Matches	600	Conservative	0	Mismatches	376	Indels	3
						Gaps	1

QY 96 CTTCAACTATCTCTCTCCGACGAAATGGACCGACGTTCTTCCGACTTCAATTT 155
Db 211 CCCCTTCAACTGSCGCGCCGACGCGCTGGAGCCGACCGACTTCTTCCGACTTCAACT 270
QY 156 CACTTCGCAATTCCTCTCTCTCTCTTGGCAACGACCAACACCGGCACTTCGATTC 215

271 CTTCCCG---CCGCCCGGACCCGCGCGGCGCTCTCTTCTCTCCGTCGCCGCT 327
Qy 216 TCTGCCACCATCTATTGATTCGTCGGAAGCCACGCTCCCATCCCATTTGATTTTACCA 275
Db 328 CGTCCCGATCTTCCCGGAAGCCGCGCGCTTGGCCCTCCCGGTCGACTTCTACAA 387
Qy 276 GGTATTAGAGAGCTCAACACATTTCTTAACCGATGAATCAAGAGAGATTGGAAGTAG 335
Db 388 GATTCTTGCTGGGAGCCACATTTCTAGGCGATGGCATTCGGAGGGCGTTGAGTCCG 447
Qy 336 GGTTCGAAACCGCGCAATTCGTTTGAGGAGACGCTTTAATCACCCGAGACAT 395
Db 448 GATAGCTAAGCCTCCTAGTATGGTACAGACAGAACTCTTGTGGACGCGCAAT 507
Qy 396 TCTTCAGACTGCTTCGGAACCTCTCTAATCTCTGCTCTAGAGAGAGTCAATGAAG 455
Db 508 GCTGAGATTCGCCATGATCTCTCAACACAGAGCTTCGCGACCGAGTACGACCG 567
Qy 456 TCTTCTGATGATGAGAGAGCTACAGTCACTGATGTTCTCTGGATTAAGTTCTG 515
Db 568 GCTTCCGAGACCGCTGATGCGGCACTCACATGATGCTGCTGGATTAAGTTCCAG 627
Qy 516 GCTCTCTGATGATGAGAGAGGCTGAGAGCTGAGATGATCTTCGGGTTGAGAGC 575
Db 628 TGTGCTGTGTGCTTCAAGAGGCTGGAGGACACACTGCTGTTAGCACTGAGAGCA 687
Qy 576 TCTGCTTAAGAGAGAGTTCCTAAGTCCGTTTAAGCAAGATGCTTTAGTTATGCGCT 635
Db 688 CTGCTTCAAGAGACGCTCAACCTAAGCGGTTGAGAGAGATGTGGTCTAGCAATGCAAT 747
Qy 636 TCGCTTCTCGATGCTCTGAGGAGATGCTATGAGCATTTGATCCACTGATTTTATTA 695
Db 748 GCTTATCTGAGCATATCAAGGATGCTATGCGACAGACCTCCAGATGATCTGCTG 807
Qy 696 TTATGATTTGTTGAGAGAGCTTTGAGCTTTTACAGAGAGAGAGAGAGAGCTTGC 755
Db 808 TTGTAGGCTCTTGAAGAGGACCTGAAGCTCTGAGAGAGATGCGAGAGCATCTTGC 867
Qy 756 ACCGATTTAAGAGAGAGTTCATGATGAGATTTGAGAGAGATCTCCGCTTATGCTT 815
Db 868 AACTGAACTGCTTACAGATTTGATGAACTTTGAGAGAGATTAACCTCGTTGTGAT 927
Qy 816 GAGGCTACTGAGCTTACCGCTTGTGATGATTAACCTCGAGAAAGACTAAATGTTAAG 875
Db 928 GAGGCTTCTGCTCTCTACTGATGAATAAATTAACCCCAAGAGGCTGCA 987
Qy 876 CGGTGCGGAGATTTTGTGCTGTTGAGAGAGTGAAGATCAGCTTTGTTGGGG 935
Db 988 AGGTGCAAAAACATATTGTGAGTGTGGCAGGGGTGTATGTCTACTGTTGAGAGG 1047
Qy 936 TTTGACCCGAGAGATTTATGATGAGGCGTTTTCGAATGACAGCTGCTGAGCAGT 995
Db 1048 ATTTCTGCTGAGGCTTCATGAATGAGGCTTTCTTGAGATGACATCAGCTGAGCAGAT 1107
Qy 996 TGATCTTTTGTAGCTACCCCAAGCAATTTCCAGAGATGATTTGAAGTTTACGAAGT 1055
Db 1108 GGAATTTCTTCTTAACACCGAATAGCATACCACTGAATGTTTGAATCTATAGTGT 1167
Qy 1056 TGCACTGCTCTGTTGGCT 1074
Db 1168 CGCACTTCCCACTTGGCT 1186

RESULT 12

US-10-600-070-184

; Sequence 184, Application US/10600070
; Publication No. US2004013950A1
; GENERAL INFORMATION:
; APPLICANT: Osteoryoung, Katherine W.
; APPLICANT: Vithe, Stanislav
; APPLICANT: Kokeharova, Olga A.
; APPLICANT: Gao, Hongbo
; TITLE OF INVENTION: Placid Division and Related Genes and Proteins, and Methods of

; TITLE OF INVENTION: Use
; FILE REFERENCE: MSU-08153
; CURRENT APPLICATION NUMBER: US/10/600,070
; CURRENT FILING DATE: 2003-06-20
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: Patent version 3.2
; SEQ ID NO 184
; LENGTH: 631
; TYPE: DNA
; ORGANISM: Prunus persica
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (21)..(21)
; OTHER INFORMATION: n is a, c, g, or t
US-10-600-070-184

Query Match 13.7%; Score 328.8; DB 7; Length 631;
Best Local Similarity 71.5%; Pred. No. 3,4e-88;
Matches 451; Conservative 0; Mismatches 168; Indels 12; Gaps 1;

Qy 913 GAGCATCAGCTCTTGTGGGGGTTTGACCCGTGAGAGTTTATGATGAGCGCTTTT 972
Db 1 GCACTTGCATTCGTGGGGGATTCATCTACGTGAATTTTCAAGACAGGCTTCTTG 60
Qy 973 GGAATGACAGCTGCTGAGCAGTTGATCTTTTGTAGCTAACCCAGCAATATTCAGCA 1032
Db 61 CATATGACTCAGCTGAGCAGTTGATTTTGTAGCTACCCCACTAATATCCCGCA 120
Qy 1033 GAGTCATTAAGTTTACGAATGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1092
Db 121 GAAAGCTTGAAGTTATGAGGATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Qy 1093 AAGCCACCTTTTACAGATGCTGATTAAGCAATTCAGCACTTCAGCAGCTTAAGTA 1152
Db 181 AAACCTCATCACTTAAGATGCTGAAGAACTTATCCAGAACTTCAGAGCTTAAGTA 240
Qy 1153 ATGCTATGAGATCTCTGCAATGTTTATGATACACGAAATATTTGGAGATGACTTC 1212
Db 241 ACAGCTTAGACATTCCTTGAACACTATTAACCAAGAAAGCAGTGAATAGACTTT 300
Qy 1213 GGTCTAAGAGGGGACTCTGACCTGCTATTAAGGCAAGATGATGAATCCGATGTGG 1272
Db 301 GCTTGAAGAGGGGACTCTGCTTCTCTGAGGAGCCTTGATGACATGCTGTTGCTGG 360
Qy 1273 TTGGGCTTGAAGCAGTGGATTCACATATATGAGATTCAGCTATTTGAGATTTTGG 1332
Db 361 TTGGGCTTGAAGCAGTATGATTAATGATTAAGAAATCATCTGTTGATGACTTTGCTTG 420
Qy 1333 GAGATTCGAATCTGATG-----ACATGATGATCTCCTGGACTATGCAAA 1380
Db 421 GAGAACTCAAGAGATGACATGACATATGACATATGATCTTCTGAGACTTTGCAAG 480
Qy 1381 TTGTGGAACCTGGTGGAGGGGTTGCTTCTGATGTTCAAGAGACACCAAGATTA 1440
Db 481 CTATTTGAGACGTGATGATGAGAGGTGATTTCCCGAGTTTGAAGACCAAGACATA 540
Qy 1441 AAATTTAACTCGGGGACTATGATGATCTATGATCTTATGATTTTGAATTTTGAAGAGT 1500
Db 541 GAGTTGAGACTGGGAGCTACTATGATGATCTTAAGTCTTGAATATCTTAAGAAAGCTG 600
Qy 1501 GAGTATGTCAGGGTTCTCTTAACTGCTG 1531
Db 601 GATGCACTAATGTTCACTTAACTGCTG 631

RESULT 13

US-10-600-070-135

; Sequence 135, Application US/10600070
; Publication No. US2004013950A1
; GENERAL INFORMATION:
; APPLICANT: Osteoryoung, Katherine W.
; APPLICANT: Vithe, Stanislav
; APPLICANT: Kokeharova, Olga A.

APPLICANT: Gao, Hongo
TITLE OF INVENTION: Placid Division and Related Genes and Proteins, and Methods of
FILE OF INVENTION: Use
FILE REFERENCE: MSU-08153
CURRENT APPLICATION NUMBER: US/10/600,070
CURRENT FILING DATE: 2003-06-20
NUMBER OF SEQ ID NOS: 206
SOFTWARE: PatentIn version 3.2
SEQ ID NO 135
LENGTH: 660
TYPE: DNA
ORGANISM: Medicago truncatula
US-10-600-070-135

Query Match 11.6%; Score 278.8; DB 7; Length 660;
Best Local Similarity 67.3%; Pred. No. 5.2e-73;
Matches 432; Conservative 0; Mismatches 192; Indels 18; Gaps 2;

81 AGGTAGCCACACACCTCTACAACTATCTGCTCCGACGAAATGGGCGGACCTTTCT 140
17 ACTTAACCGCTCTCCATCTCCGCGCGCTCTCCGACAGTAATGGGCGGACGACTCAT 76
141 CTCGCACTCAATTTGACCGGATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 200
77 TTCCGATTTCCATTTCTCGGCGACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 136
201 CGCACT 260
137 AGTCACTCTC-----ACTCTTTCTTAACCTCTCTCTCTCTCTCTCTCTCTCTCTCT 190
261 CATTGATTTCTACAGGATTTAGAGCTCAACACATTTCTTAACCGATGAAATCAAG 320
191 TCTCGACCTGTCAAAATCTCGGCGCGCAAAACGCTTTCTCGATGATGATTCGAG 250
321 AGCATTCGAAGCTAGGCTTTCGAAACCGCGCAATTCGCTTCTCTCTCTCTCTCTCT 380
251 AGCTTATGAAGGAATTTCTGAAGCTCTCTCAATGCTTCTCAATGAAGCTTTGAT 310
381 CAGCCGAGACAGATTTCTCAAGCTGCTTGCAGAACTCTGCTAATCTCTGCTTGAAG 440
311 TGTGCTGCTGAGATTTCTCAAGCTGCTTGTGAACCTCAAGCTGCTCTCTTGAAG 370
441 AAGATCAATGAAGCTTCTTGTGATGAA-----GAACTCAAGTCAATCAC 488
371 AAGATATATCAAAAGCTGCTGACATGAAGACGAAGATGAGATCTTCAATCTCAC 430
489 TGAATGCTCTTGGGATTAAGTCTCTGAGGCTCTCTGATTTGAAGAGTGTGAGAC 548
431 TAAATCCCTTTCGACAAAGTCTGAGCTCTGCTGCTTGAAGAGTGTGAGAGAC 490
549 TGAGATAGTTCTTGGGTTGTGAGGCTCTGCTTAAGAGAGTGTGCTTAAGTCTTGA 608
491 GAGTGTGCTTGGATTTGAGGGGCTTACTGAGAGAGAGTTAACCAATGTTTAA 550
609 GCAAGATGTTGTTTGAATGATGAGGCTTCTGATGCTGAGAGATGCTATGAC 668
551 GCAAGATGTTGTTGAGGCTGATGAGGCTTCAATATGTTGAGCTTCAAGGATGCTATG 610
669 ATTGATCCACCTGATTTTATACGTGTTATGAGATTTGTTGA 710
611 TTGTCCCGCCAGATTTTATGTTGCTTGTGAGAGTCTGGA 652

RESULT 14
US-10-424-599-35059
Sequence 35059, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Gao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 35059
LENGTH: 1039
TYPE: DNA
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_131660C.1
US-10-424-599-35059

Query Match 9.6%; Score 231.6; DB 7; Length 1039;
Best Local Similarity 64.4%; Pred. No. 1.5e-58;
Matches 380; Conservative 0; Mismatches 204; Indels 6; Gaps 2;

1823 TTGCTGATATGTTAAAGAGGCAAGTGAAGATCTTACGCTGCTGTGCAATTGAC 1882
50 TTACTGAAGAAATTAACATGCAAGTGTGACATCATGTGTGTGTAAATGGAC 109
1883 TGATTTCACTGTTCAGCAAGATTTTCTTAAAGCA-----GCTCATCTTTCAAGCA 1939
110 TGGTACTTGTGTGTTTAAAGTTTTAACTTCACTTGAAGATGCTACCATTTCTGTA 169
1940 AGGATATGTTTCTTCTATGAAATCTGATGCTGCTACCATAGAGGTCAGTCAAGCTGAC 1999
170 AATGACGTGTTCCGCAATGTGTGATGATCATTAATTAAGTTTCTTGGAGATGAAG 229
2000 ATTCAAGG---CACTTCCCAAGATGATGCTAGACCTGACAGAGAAATTAATCAAGT 2056
230 AAAAGTGAAGCAGCTACCAAAATGATGCAAGGTTGCCGAGGCTTTAGTCCCAAGT 289
2057 GGCAGAAATTAAGTCTGTGCTTTTGGGCGCGATCACCGATGAATGAATTAACAGAG 2116
230 GGCAGAGTAAATTCGAAAGCTTTTGAACCTGACATGCTTGGAGGTTGATGAGG 349
2117 TTTTGAATGAGGCAATGCTGAAATTTGAACTGACAGACAGCTGAACTGCGACGTTG 2176
350 TTTTGAAGTGAAGATGTTGAAGATGAGACAGATCGTGAAGTGAATGAGAGCGTG 409
2177 GGTGCTTATGATTAATACCTGTTGAACATCTCTGTTGACAGTGTGACAGTCTGACAG 2236
410 GTTGTCTTATGATGATCACTTGTGAGGACCTCAACATGAAATGATGATCAATCACGA 469
2237 ATGAACCGCTGCTGTGAGAGCACTGAGAGTCTGCTGCTATCTGATTTGG 2296
470 ATGGCGCTGCTGAGTGTGAAACACTTGAAGAGTCTTACTACCTGATGCCGTA 529
2297 TTCAATCCAGAAACAAATGCTATGATGACAGACCTACACAAAGATGAGATTTCT 2356
530 GTCAATCCGCAACAGATCTTCTAATAGCAGACCTACACAAAGATGATGATGCTT 589
2357 GGTCAAGTCAAGGTGAGAAATCACTGAAGGCTGTCTTTCATCATTA 2406
590 TTACAGGCGCAGAGTGAATTTGTGAAGAGAGCTGTCTTGAATCTTAA 639

RESULT 15
US-10-600-070-143
Sequence 143, Application US/10600070
Publication No. US20040139500A1
GENERAL INFORMATION:
APPLICANT: Ostryoung, Katherine W.
APPLICANT: Vitha, Stanislaw
APPLICANT: Koksharova, Olga A.
TITLE OF INVENTION: Placid Division and Related Genes and Proteins, and Methods of
FILE REFERENCE: MSU-08153
CURRENT APPLICATION NUMBER: US/10/600,070
CURRENT FILING DATE: 2003-06-20
NUMBER OF SEQ ID NOS: 206

; SOFTWARE: Patentin version 3.2
; SEQ ID NO 143
; LENGTH: 537
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-10-600-070-143

Query Match 9.3%; Score 224; DB 7; Length 537;
Best Local Similarity 63.6%; Pred. No. 1.9e-56;
Matches 341; Conservative 0; Mismatches 195; Indels 0; Gaps 0;

```
QY 555 AGTTCTTCGGGTTGAGAGGCTCTGTGTTAAGAGAGGTTGCTTAAGCTTTAAGCAGA 614
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 2 AGTGTTCGCAATTGAGAGGACCTTACTGAGAGCCGCCCAAGCGGTTCAAGCAGGA 61
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 615 TGTGGTTTAAATTATGAGGCGCTTCCGTTCTGATGCTCGAGAGATGCTATGCAATTGGA 674
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 62 TGTGTGCTGGCAATGCGCGCTGCTGATGTGATCTATCAAGAGAGCGCAATGCGGCTAG 121
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 675 TCCACCTGATTTTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 734
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 122 CCTCCAGATGTAATCCGCTGCTGAGAGGCTTGAAGGCTCTCAAGCTTTTGACAGA 181
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 735 GGAAGAGCAATAGACCTTGCACCGGATTTACGTGACAAATTGATGAGACTTTGGAAGA 794
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 182 GGATGGGGCAATCAATCTCGCACTGGTTTGTCTCACAATTTGATGAAACTCTGAGGA 241
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 795 GATCACTCCGCGTTTATGCTTGTGAGCTACTTGGCTTACCGCTTGGTATGATTAAGCTGC 854
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 242 TATCACTCCTCGTTGTGTTTGAAGCTTTTCCCTTCTCTTGTATGAAAAACATCAGAA 301
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 855 GAAAAGACTAAATGTTTAAAGCGGTGTCGGAATAATTTGAGTCTGTTGAGAGAGTGG 914
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 302 TGAACACCAAGAGGCTCTGCTGTGTGGAACAATTTTGTGAGAGTGTGGCAGAGAGG 361
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 915 AGCATCAGCTCTTGTGTTGGGGTTTGAACCGGTGAGAGATTTATGATGAGGCGTTTACG 974
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 362 TATTGGTACTGTTGAGAGGATTTTCGGTGAAGCTTACATGAATGAAGCTTCTGCA 421
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 975 AATGACAGCTGCTGAGAGAGTGTATCTTTTGTAGCTACCCCAAGCAATATTCACAGAGA 1034
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 422 GATGACATGCGCGAGAGAGATGATTTCTTCAAAAACACCGAATAGCATACCGCTGA 481
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1035 GTCAATTTGAAGTTTGAAGATGCACTTGCTGCTGCTCAAGCTTTTATTTGTA 1090
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 482 ATGATTGAAATCTATAGCGTGCACTTGCAATGTTGCTCAAGCAATTTGTAAGTA 537
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Search completed: December 10, 2005, 18:15:31
Job time : 1850.5 secs

SEQ ID NO 12844
LENGTH: 57280
TYPE: DNA
ORGANISM: Human
US-09-949-016-12844

Query Match 2.1%; Score 50.2; DB 3; Length 57280;
Best Local Similarity 53.9%; Pred. No. 0.0059;
Matches 103; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

85 AGCCACAACCTCTACAACTATGCTCCGACGAATGGCGGACCGTCTCTCC 144
13101 AGCACCATCACCACCACTGCTCCACCACTACCACTTACCTATGAGCTCC 13042
145 GACTTCAATTTACCTCCGATTCCTCTCTCTCTCCGACCGGACCAACCGCC 204
13041 ACCATCATATATCCATATACGCTCCACCAACCGCTTACCATACCAACCAACC 12982
205 ACTCTGCTCTCTGCGACCACTATGATGCTCCGAAAGCGCAGTCCCATCCCAT 264
12981 ACAACCACTACTCATATGATGATCTTCCATCAACATTCATCACCCTCCACCAAC 12922
265 GATTCTACCA 275
12921 CATCATCACA 12911

RESULT 5
US-09-949-016-12846/c
Sequence 12846, Application US/09949016
Patent No. 6812339

GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 12846
LENGTH: 57280
TYPE: DNA
ORGANISM: Human
US-09-949-016-12846

Query Match 2.1%; Score 50.2; DB 3; Length 57280;
Best Local Similarity 53.9%; Pred. No. 0.0059;
Matches 103; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

85 AGCCACAACCTCTACAACTATGCTCCGACGAATGGCGGACCGTCTCTCC 144
13101 AGCACCATCACCACCACTGCTCCACCACTACCACTTACCTATGAGCTCC 13042
145 GACTTCAATTTACCTCCGATTCCTCTCTCTCTCCGACCGGACCAACCGCC 204
13041 ACCATCATATATCCATATACGCTCCACCAACCGCTTACCATACCAACCAACC 12982
205 ACTCTGCTCTCTGCGACCACTATGATGCTCCGAAAGCGCAGTCCCATCCCAT 264
12981 ACAACCACTACTCATATGATGATCTTCCATCAACATTCATCACCCTCCACCAAC 12922
265 GATTCTACCA 275
12921 CATCATCACA 12911

RESULT 6
US-09-949-016-13542/c
Sequence 13542, Application US/09949016
Patent No. 6812339

GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13542
LENGTH: 57280
TYPE: DNA
ORGANISM: Human
US-09-949-016-13542

Query Match 2.1%; Score 50.2; DB 3; Length 57280;
Best Local Similarity 53.9%; Pred. No. 0.0059;
Matches 103; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

85 AGCCACAACCTCTACAACTATGCTCCGACGAATGGCGGACCGTCTCTCC 144
13101 AGCACCATCACCACCACTGCTCCACCACTACCACTTACCTATGAGCTCC 13042
145 GACTTCAATTTACCTCCGATTCCTCTCTCTCTCCGACCGGACCAACCGCC 204
13041 ACCATCATATATCCATATACGCTCCACCAACCGCTTACCATACCAACCAACC 12982
205 ACTCTGCTCTCTGCGACCACTATGATGCTCCGAAAGCGCAGTCCCATCCCAT 264
12981 ACAACCACTACTCATATGATGATCTTCCATCAACATTCATCACCCTCCACCAAC 12922
265 GATTCTACCA 275
12921 CATCATCACA 12911

RESULT 7
US-09-949-016-13543/c
Sequence 13543, Application US/09949016
Patent No. 6812339

GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 13543
LENGTH: 57280
TYPE: DNA
ORGANISM: Human
US-09-949-016-13543

Query Match 2.1%; Score 50.2; DB 3; Length 57280;
Best Local Similarity 53.9%; Pred. No. 0.0059;

	Matches	103;	Conservative	0;	Mismatches	88;	Indels	0;	Gaps	0;
QY	85	AGCCACAACACCTTACAACTATCTGTCGCCAGCAATAGGCGCACCGCTCTTCTTCC								144
Db	13101	AGCAACCATACACACCAACATGGCTCTCACACCATATACCACTTACACCATTTATGAGCTTC								13042
QY	145	GACTTCATTTTCAACCTCGATTCTCTCTCTCTCTTGGCCACGCGACCAACCAACGGCC								204
Db	13041	ACCATTAATATCCCTATCAGGTTCACACCAACCGCTTTCAACATATACACACCAACCAAC								12982
QY	205	ACTCTGCTCTCTGCGACCATTTATGATGTCGCCAAGCCACGTGCCATTCGCCATT								264
Db	12981	ACAACCGACCACTCACTATCATGACTTCACACCATCAACAATTACCATCACTCCACACAAC								12932
QY	265	GATTTTCACCA	275							
Db	12921	CATCATACGA	12911							

```

RESULT 8
US-09-949-016-13544/C
; Sequence 13544, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 13544
; LENGTH: 57280
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13544

Query Match      2.1%; Score 50.2; DB 3; Length 57280;
Best Local Similarity 53.9%; Pred. No. 0.0059;
Matches 103; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY      85  AGCCACAACCTCTACAACTATCTGCTCCGCCAGCAAAATGGCCGACCGCTCTTCTCC 144
Db      13101 AGCAACCATCACCAACCAACATCTGCTCCACCAACATCACCACTTACCAATTGAGCTCC 13042

QY      145  GACCTCAATTTTCAACCTCGATTCCTGCTCCTCTTGGCCACGCGCAACCAACGCGC 204
Db      13041 ACCATCAATATCTCTATACGCTCCACACCAACCGTTTACCATACCAACCAACGCGC 12982

QY      205  ACTCTGCTCTCTCTCCCAACCATCTATTGATGCTGCGGAAGCGCACGCTCCCATGCCCAT 264
Db      12981 ACAACCAACACGTCACGATCATGATACCTTCACCATCAATTCACATCACTCCACCAAC 12922

QY      265  GATTTCTACCA 275
Db      12921 CATCATCACGA 12911

RESULT 9
US-09-949-016-13545/C
; Sequence 13545, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

```

```

; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 13545
; LENGTH: 57280
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13545

Query Match      2.1%; Score 50.2; DB 3; Length 57280;
Best Local Similarity 53.3%; Pred. 0.0059;
Matches 103; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY      85  AGCCACACACCTCTTACACTATATCTGCTCGGCACGAAATGGCGCAGCCGTTCTCTTC 144
Db       13101 AGGACCCATTCACACCAACACTGGCTCCACACCATCACCACCTTCACTTATGAGCTTC 1304283

QY      145  GACTTCATTTGACCTTCGATTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 204
Db       13041 ACCATCAATATCCCTATCAGCTTCACACACACACACCGTCTTTACATCATTACACACACAC 1298283

QY      205  ACTCTCGCTCTCTGTGCACCATCTATGTATCGTCCGAAAGCCACAGTCCTCCATTCGCCATT 264
Db       12981 ACAACACACACTCATCTATCATGTACTTCACACATCAATCAATTACATCACTCCACACGACAC 1292222

QY      265  GATTTTACCA 275
Db       12921 CATCATCAGA 12911

```

```

RESULT 10
US-09-949-016-14633/C
; Sequence 14633, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14633
; LENGTH: 57280
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14633

Query Match      2.1%; Score 50.2; DB 3; Length 57280;
Best Local Similarity 53.9%; Pred. No. 0.0059;
Matches 103; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

Oy      145 GACTTCAATTTCACCTCCGATTCCTCCTCTCTCTCTCTCTTGCCACCGCCACACACCAGGCC   204
|||
Db      13101 AGCACCATCACACACACCACTAAGTCCCTCCACACACATCACACACTTACCACTTTAGAGCTCC   13042
|||
Oy      85 AGCCACACACCTCTACAACAATGTGCTCCGGCAGACAATAGGCGCAGCGCTCTCTCTCC   144
|||

```

Db 13041 ACCATCAATATCCCTATATCAAGTCCACCAACCGCTTTCACATACCAACCAACACC 12982
QY 205 ACTCTGCTCTCTGCGCAACCATATATGATCGTCCGGAAGCGACGCCATCCCATTT 264
Db 12981 ACAACCACTCACTACTATATGACTTCCACCATCAACATTAACATACCTCCACCAAC 12922
QY 265 GATTTCTACCA 275
Db 12921 CATCATCACGA 12911

RESULT 11
US-09-949-016-14634/c
Sequence 14634, Application US/09949016
Patent No. 6812339

GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: CLO01307
CURRENT FILING DATE: 2000-04-14
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 14634
LENGTH: 57280
TYPE: DNA
ORGANISM: Human
US-09-949-016-14634

Query Match 2.1%; Score 50.2; DB 3; Length 57280;
Best Local Similarity 53.9%; Pred. No. 0.0059;
Matches 103; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 85 AGCCACAACCTCTACAACTATGCTCCGCCAGCAATGGCGCGATCTTCTCTCC 144
Db 13101 AGCACATACACCAACCAACGCTCCACCAACCATACCACTTATGAGCTCC 13042
QY 145 GACTTCAATTTCACTCCGATTCCTCTCTCTCTCTCTGCGACGCGACCAACCGCC 204
Db 13041 ACCATCAATATCCCTATATGACTTCCACCATCAACATTAACATACCTCCACCAAC 12922
QY 205 ACTCTGCTCTCTGCGCAACCATATATGATCGTCCGGAAGCGACGCCATCCCATTT 264
Db 12981 ACAACCACTCACTACTATATGACTTCCACCATCAACATTAACATACCTCCACCAAC 12922
QY 265 GATTTCTACCA 275
Db 12921 CATCATCACGA 12911

RESULT 12
US-09-949-016-14635/c
Sequence 14635, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: CLO01307
CURRENT FILING DATE: 2000-04-14
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 14635
LENGTH: 57280
TYPE: DNA
ORGANISM: Human
US-09-949-016-14635

PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 14635
LENGTH: 57280
TYPE: DNA
ORGANISM: Human
US-09-949-016-14635

Query Match 2.1%; Score 50.2; DB 3; Length 57280;
Best Local Similarity 53.9%; Pred. No. 0.0059;
Matches 103; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 85 AGCCACAACCTCTACAACTATGCTCCGCCAGCAATGGCGCGATCTTCTCTCC 144
Db 13101 AGCACATACACCAACCAACGCTCCACCAACCATACCACTTATGAGCTCC 13042
QY 145 GACTTCAATTTCACTCCGATTCCTCTCTCTCTCTCTGCGACGCGACCAACCGCC 204
Db 13041 ACCATCAATATCCCTATATGACTTCCACCATCAACATTAACATACCTCCACCAAC 12922
QY 205 ACTCTGCTCTCTGCGCAACCATATATGATCGTCCGGAAGCGACGCCATCCCATTT 264
Db 12981 ACAACCACTCACTACTATATGACTTCCACCATCAACATTAACATACCTCCACCAAC 12922
QY 265 GATTTCTACCA 275
Db 12921 CATCATCACGA 12911

RESULT 13
US-09-949-016-14636/c
Sequence 14636, Application US/09949016
Patent No. 6812339

GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: CLO01307
CURRENT FILING DATE: 2000-04-14
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 14636
LENGTH: 57280
TYPE: DNA
ORGANISM: Human
US-09-949-016-14636

Query Match 2.1%; Score 50.2; DB 3; Length 57280;
Best Local Similarity 53.9%; Pred. No. 0.0059;
Matches 103; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 85 AGCCACAACCTCTACAACTATGCTCCGCCAGCAATGGCGCGATCTTCTCTCC 144
Db 13101 AGCACATACACCAACCAACGCTCCACCAACCATACCACTTATGAGCTCC 13042
QY 145 GACTTCAATTTCACTCCGATTCCTCTCTCTCTCTCTGCGACGCGACCAACCGCC 204
Db 13041 ACCATCAATATCCCTATATGACTTCCACCATCAACATTAACATACCTCCACCAAC 12922
QY 205 ACTCTGCTCTCTGCGCAACCATATATGATCGTCCGGAAGCGACGCCATCCCATTT 264
Db 12981 ACAACCACTCACTACTATATGACTTCCACCATCAACATTAACATACCTCCACCAAC 12922
QY 265 GATTTCTACCA 275

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 10, 2005, 16:57:24 ; Search time 11911.1 Seconds
(without alignments)
11482.131 Million cell updates/sec

Title: US-10-600-070B-1

Perfect score: 2406
Sequence: 1 atggaagctctgagtcacgt.....gtctgtcttcgacataa 2406

Scoring table: IDENTITY NUC
Gapop 10'-0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*
1: gb_da:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pac:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_sts:*
11: gb_sy:*
12: gb_un:*
13: gb_vl:*
14: gb_htg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2402.8	99.9	2438	AY221469	Arabidopsis
2	2399.6	99.7	2436	AY150490	Arabidopsis
3	2399.6	99.7	2637	AY091075	Arabidopsis
4	1724	71.7	3668	AY221468	Arabidopsis
5	1720.8	71.5	3664	AY221467	Arabidopsis
6	1717.6	71.4	85791	AB016888	Arabidopsis
7	243.6	10.1	133779	AC158210	Medicago
8	180.8	7.5	110000	AP008208_011	Continuation (12 o
9	180.8	7.5	150462	AP004885	Oryza sat
10	165.4	6.9	550	DQ022571	Beta vulg
11	163.4	6.8	184752	AC155597	Zea mays
12	144.2	6.0	117818	AC160013	Medicago
13	62	2.6	221756	AC094893	Rattus no
14	62	2.6	232974	AC114193	Rattus no
15	60.8	2.5	200412	AC115666	Rattus no
16	60.8	2.5	253094	AC113774	Rattus no
17	60.2	2.5	232373	AC113774	Rattus no
18	59.6	2.5	193341	AC134500	Rattus no

19	59.2	2.5	96573	14	AC136672	AC136672 Rattus no
20	59.2	2.5	175351	14	AC117900	AC117900 Rattus no
21	59.2	2.5	233591	14	AC121387	AC121387 Rattus no
22	59.2	2.5	243553	14	AC096320	AC096320 Rattus no
23	59.2	2.5	247341	14	AC099178	AC099178 Rattus no
24	59.2	2.5	312143	14	AC123509	AC123509 Rattus no
25	59.2	2.5	346208	14	AC128290	AC128290 Rattus no
26	57.6	2.4	7218	6	166494	166494 Sequence 14
27	57.6	2.4	256578	14	AC099432	AC099432 Rattus no
28	57.6	2.4	313096	14	AC118179	AC118179 Rattus no
29	57.2	2.4	98521	9	AC094507	AC094507 Rattus no
30	57	2.4	263036	14	AC105597	AC105597 Rattus no
31	56.8	2.4	110000	14	CR954204_0	CR954204 Ostrycoc
32	56.4	2.3	283355	14	AC104401	AC104401 Rattus no
33	56	2.3	143976	14	AC141543	AC141543 Rattus no
34	56	2.3	246199	14	AC119558	AC119558 Rattus no
35	56	2.3	252996	14	AC130153	AC130153 Rattus no
36	56	2.3	253772	14	AC109777	AC109777 Rattus no
37	55.6	2.3	242516	14	AC106523	AC106523 Rattus no
38	54.6	2.3	165849	14	AC114251	AC114251 Rattus no
39	54.6	2.3	231716	9	AC135409	AC135409 Rattus no
40	54.6	2.3	234812	14	AC111482	AC111482 Rattus no
41	54	2.2	110000	14	AC129796_0	AC129796 Rattus no
42	54	2.2	240800	14	AC095964	AC095964 Rattus no
43	54	2.2	274697	14	AC098240	AC098240 Rattus no
44	53.8	2.2	205796	14	AL713891	AL713891 Mus muscu
45	53.8	2.2	215524	9	AL603913	AL603913 Mouse DNA

ALIGNMENTS

RESULT 1
AY221469 2438 bp mRNA linear PLN 05-AUG-2003
LOCUS Arabidopsis thaliana division protein (ARCE) mRNA, complete cds;
DEFINITION nuclear gene for chloroplast product.
ACCESSION AY221469
VERSION AY221469.1 GI:33436353
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 2438)
Vitha,S., Froehlich,J.E., Koksharova,O., Pyke,K.A., Van Eyrp,H. and Oosteryoung,K.W.
ARC6 is a J-Domain Plastid Division Protein and an Evolutionary Descendant of the Cyanobacterial Cell Division Protein Ftn2
JOURNAL Plant Cell 15 (8), 1918-1933 (2003)
PUBMED 12897262
REFERENCE 2 (bases 1 to 2438)
Vitha,S., Koksharova,O., van Eyrp,H., Froehlich,J.E. and Oosteryoung,K.W.
Direct Submission
Submitted (17-JAN-2003) Department of Plant Biology, Michigan State University, 166 Plant Biology Building, East Lansing, MI 48824, USA
FEATURES
source
location/Qualifiers
1..2438
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/ecotype="Wassilewskija"
1..2438
/gene="ARC6"
/note="synonym: At5g42480"
2..2407
/gene="ARC6"
/codon_start=1
/product="division protein"
/protein_id="AAQ18646.1"
/db_xref="GI:33436354"

```
/translation="MEALSHVIGLSFOLCRLEPATTYKLRSHNTSTTICSAKMAD
RLISDFNPSTSSSSSPATATATATATVLSLPISIDRPERHPIDPEQVIGAOHPFT
DJIRAPFARVSKPPOFSDDALISROLOAECITLSNPSRREPEVGLDEBEGAT
VITDPWDKVPALCVLOBGEETRIIVRVEGALKERLPGSDVULWALAFVPLDS
RDMAALDPDPFITGYEFVEBEKILQSEGNASSLPDRAQIDETLEBITTRYVLELIG
LPIDUDYAKRNLGSLGVRNLSVGGGASVALVGLIREFPNKEATLRMAEVOVL
FVATPSP1PAPSEFVEYVALVAQAF1GKKPHILOADQPOLOQAKWAME1PAM
LYDRNNMEIDFGLERGLCALIGKVECKRMWLGJDESDQYRNPAIVEVELENSND
DNDLPGLCKLETWLAGVVEPRPRDKKFKLDGYDDPMTVSYLERVEVQSGPL
AAATMARIGAEHYKASAMOLOKVPPSRYTDRNSAPKPOVETPFSDVGNVGRD
GEBGVFIABAVREPMFETNDYAIRAVSESSEVDETVEVMSVAMLEKASVILAAAV
AIGLISLPSQKIFLKSSSFQKDWSSMESDVATIGSVRADSEALPRMDAKIENI
VSKWQK1KSLAFEDRHER1EM1PEVLDRMLKIWDRAEATQALQVYDYLTKLSVDS
VTVASDGTBALVEATLEESAC1SDLVHPENNATDVRTYTRIEVFMKSGMKITGSGV
LAS"
misc_difference 474
/gene="ARCG"
/feature="results in Gly to Glu substitution"
/misc_difference 1831
/gene="ARCG"
/feature="results in Glu to Asp substitution"
/replace="t"
misc_difference 1831
/replace="t"
ORIGIN
Query Match 99.9%; Score 2402.8; DB 15; Length 2438;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2404; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 661 GCTATGCGATTGATCCACCTGATTTTATTAATCGTTATAGTTTGTGAGAACCTTGG 720
Db 662 GCTATGCGATTGATCCACCTGATTTTATTAATCGTTATAGTTTGTGAGAACCTTGG 721
QY 721 AAGCTTTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
Db 722 AAGCTTTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 781
QY 781 AAGCTTTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
Db 782 AAGCTTTTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 841
QY 841 GATGATTAACGCTCGAAGAAAGCTTAAATGTTTAAAGCGGTGCGGAATATTTTGCT 900
Db 842 GATGATTAACGCTCGAAGAAAGCTTAAATGTTTAAAGCGGTGCGGAATATTTTGCT 901
QY 901 GTTGGAGAGAGTGGAGATCACTCTTGTGGGGGTTTGAACCCGTGAGAAATTATGAT 960
Db 902 GTTGGAGAGAGTGGAGATCACTCTTGTGGGGGTTTGAACCCGTGAGAAATTATGAT 961
QY 961 GAGGCGTTTTTACGAATGACAGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
Db 962 GAGGCGTTTTTACGAATGACAGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1021
QY 1021 AATATTCACAGACAGATCATTTTGAAGTTTACGAAGTTGCACTTGTGTGCTCAAGCT 1080
Db 1022 AATATTCACAGACAGATCATTTTGAAGTTTACGAAGTTGCACTTGTGTGCTCAAGCT 1081
QY 1081 TTTATTTGTAAAGAGACACACCTTTTACAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
Db 1082 TTTATTTGTAAAGAGACACACCTTTTACAGAGAGAGAGAGAGAGAGAGAGAGAG 1141
QY 1141 CAGGCTAAGGTAATGAGTATGAGATTCGCGATGTTGATGATACAGGAATTAATGG 1200
Db 1142 CAGGCTAAGGTAATGAGTATGAGATTCGCGATGTTGATGATACAGGAATTAATGG 1201
QY 1201 GAGATGAGACTTCGCTCTAGAAAGGGAATCTGTGCACTGCTTATAGGCAAGTTGAT 1260
Db 1202 GAGATGAGACTTCGCTCTAGAAAGGGAATCTGTGCACTGCTTATAGGCAAGTTGAT 1261
QY 1261 TGCCGTATGTTGGGCTTAAAGACATGAGAGATTCACATATATGAAATCCAGTATTTG 1320
Db 1262 TGCCGTATGTTGGGCTTAAAGACATGAGAGATTCACATATATGAAATCCAGTATTTG 1321
QY 1321 GAGTTGTTTGGAGAAATCAATCGTATGACAAATGATATCTCCCTGGAATATGCAAA 1380
Db 1322 GAGTTGTTTGGAGAAATCAATCGTATGACAAATGATATCTCCCTGGAATATGCAAA 1381
QY 1381 TGTGTTGAAACCTGTTGGAGGGGTTGTCTTCTTCTAGGTTCAAGACACCAAGATTA 1440
Db 1382 TGTGTTGAAACCTGTTGGAGGGGTTGTCTTCTTCTAGGTTCAAGACACCAAGATTA 1441
QY 1441 AATTTTAACTCGGGGACATATATGATGATCCATATGTTTGTGTTTGTGTTTGTGTT 1500
Db 1442 AATTTTAACTCGGGGACATATATGATGATCCATATGTTTGTGTTTGTGTTTGTGTT 1501
QY 1501 GAGGTAGTTCAAGGTTCTCTTTAGTGTGCTGCAACTATGAGCAAGATTGAGGCGAG 1560
Db 1502 GAGGTAGTTCAAGGTTCTCTTTAGTGTGCTGCAACTATGAGCAAGATTGAGGCGAG 1561
QY 1561 CATGTGAAGCTAGTGTATGACAGCACTGACAGAAAGTTTCTTCCGCTATATACAT 1620
Db 1562 CATGTGAAGCTAGTGTATGACAGCACTGACAGAAAGTTTCTTCCGCTATATACAT 1621
QY 1621 AGAATCCGCTGACACCAAGAGATGACAGAGACAGTGTATATGATATCTGTTGGT 1680
Db 1622 AGAATCCGCTGACACCAAGAGATGACAGAGACAGTGTATATGATATCTGTTGGT 1681
QY 1681 AACATGTAGGCGCTGATGAGCTGAGCTGTTTATTTATGCAAGAGCTGTAAACCTCT 1740
Db 1682 AACATGTAGGCGCTGATGAGCTGAGCTGTTTATTTATGCAAGAGCTGTAAACCTCT 1741
QY 1741 GAAAACTTTGAAACTATATGATTAATGCAATTCGAGCTGGGGTCTCAAGAGATGAGCTT 1800
```

Db 1742 GAAACCTTGAACCAATGATTAATCAATTCAGAGCTGGGCTCCAGAGATGACGCTTAT 1801
Qy 1801 GAACTACTGTTGAATATGCTGCTGATATGTTAAAGAGCAGTGTGAAGATCCTA 1860
Db 1802 GAACTACTGTTGAATATGCTGCTGATATGTTAAAGAGCAGTGTGAAGATCCTA 1861
Qy 1861 GCTGCTGTTGGCAATTTGAGCTGATTTCACTGTTCAAGCCAGAGATTTTCTTAAAGC 1920
Db 1862 GCTGCTGTTGGCAATTTGAGCTGATTTCACTGTTCAAGCCAGAGATTTTCTTAAAGC 1921
Qy 1921 AGCTCATCTTTTCAACGCAAGATATGTTTCTTATGAGATCTGATGTGCTTACCTA 1980
Db 1922 AGCTCATCTTTTCAACGCAAGATATGTTTCTTATGAGATCTGATGTGCTTACCTA 1981
Qy 1981 GGGTCAGTCAGAGCTGAGATTTGAGATTTGAGATTTGAGATTTGAGATTTGAGATTT 2040
Db 1982 GGGTCAGTCAGAGCTGAGATTTGAGATTTGAGATTTGAGATTTGAGATTTGAGATTT 2041
Qy 2041 AATATAGTATCCAGAGTGTGAGAGATTAAGTCTGCTGCTGCTGCTGCTGCTGCTGCT 2100
Db 2042 AATATAGTATCCAGAGTGTGAGAGATTAAGTCTGCTGCTGCTGCTGCTGCTGCTGCT 2101
Qy 2101 GAAATGTTACCAAGCTTTTGGATTTGAGATTTGAGATTTGAGATTTGAGATTTGAG 2160
Db 2102 GAAATGTTACCAAGCTTTTGGATTTGAGATTTGAGATTTGAGATTTGAGATTTGAG 2161
Qy 2161 GAACTGCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2220
Db 2162 GAACTGCGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2221
Qy 2221 GTGACAGTCTCAGCAGATGAGAACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2280
Db 2222 GTGACAGTCTCAGCAGATGAGAACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2281
Qy 2281 TGTCTATCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2340
Db 2282 TGTCTATCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2341
Qy 2341 AGATACGAGATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2400
Db 2342 AGATACGAGATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2401
Qy 2401 TCATTA 2406
Db 2402 TCATTA 2407

RESULT 2
AY150490 2436 bp mRNA linear PLN 23-SEP-2002
LOCUS Arabidopsis thaliana unknown protein (At5g42480) mRNA, complete
DEFINITION cde.
ACCESSION AY150490
VERSION AY150490.1 GI:33297711
KEYWORDS FLI CDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi-
s 1 (bases 1 to 2436)
REFERENCE 1 Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M.,
Quach,H.L., Tang,C., Toriumi,M., Wallender,E.K., Wong,C., Wu,H.C.,
Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y.,
Ishida,J., Jones,T., Kamiya,A., Kawai,T., Kim,C.J., Narusaka,M.,
Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shim,P.,
Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W.,
Ecker,J.R. and Theologis,A.
Arabidopsis Open Reading Frame (ORF) Clones
TITLE Unpublished
JOURNAL 2 (bases 1 to 2436)
REFERENCE Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M.,

Quach,H.L., Tang,C.C., Toriumi,M., Wallender,E.K., Wong,C.,
Wu,H.C., Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R.,
Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Kawai,T., Kim,C.J.,
Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M.,
Shim,P., Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K.,
Davis,R.W., Ecker,J.R. and Theologis,A.
Direct Submission
Submitted (13-SEP-2002) Plant Gene Expression Center, 800 Buchanan
Street, Albany, CA 94710, USA
The RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN
Arabidopsis Full-length cDNA') : Seki,M., Narusaka,M., Ishida,J.,
Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
Hayashizaki,Y. and Shinozaki,K.
The Salk, Stanford, PGEC (SSP) Consortium members constructed and
sequenced the pUNI (ORF) clones using the RAFL cDNAs: Yamada,K.,
Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M.,
Quach,H.L., Tang,C.C., Toriumi,M., Wallender,E.K., Wong,C.,
Wu,H.C., Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R., Jones,T., Kim,C.J.,
Nguyen,M., Palm,C.J., Shim,P., Southwick,A., Tripp,M.G., Wu,T.,
Davis,R.W., Ecker,J.R. and Theologis,A.
Yamada,K. (SSP/PGEC) and Seki,M. (RIKEN GSC) contributed equally
to this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP
/PGEC) contributed equally to this work as PIs.
Annotation is based on the January 2002 version of the Arabidopsis
genome submitted to GenBank.
FEATURES
source
1..2436
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="5"
/clone="U19395"
/ecotype="Columbia"
/note="This clone is in pUNI 51."
1..2436
/gene="At5g42480"
1..2406
/gene="At5g42480"
/codon_start=1
/evidence=experimental
/product="unknown protein"
/protein_id="AA12907.1"
/db_xref="GI:23297712"
/translation="MEALSHVIGTGLSPFOCLRPATTKLRSHNTSTTCSAKMAD
RLSDFNFTSDSSSPATATTTATTVSPSPSIDRPRRHVPIPIDFYQVIGAOQHPLT
DGIRRAFEARVSRPQGFSDDALISRQILOAACETISNPSRREYNEGLLDEEAT
VITDVPMDKVGALCVLOEGGEIEIVRGEALLKBLTPSPFKQDVVLMALPAEDVS
RDMALDLPDPIITGVEFEERALKILDEGASSLSAPRAQIDITELIETPTRYUEILG
LPIGDVTAARKLNGLSGVRITLNSVGGGASALVGCITREKRFMAEATLRTAAEVDVL
FVATPSNIPAESFEVEVALALVAQATIGKPHILOADKQFOLOQAQWAMEIPM
LYDTRNWEIDFGLERGLCALIGKDECEMMLGLDSEDSQYRPAIVEFLVNSND
DNDLPGLCLLETWLAGVVEPRPDKDKFGDYDDPMVLSYERVAVVGGSL
AAAAAARVIAEAVKASAMQALQKVPSPRYTDRNSAPKQVOTFVSPVGNVGGD
GECQVTAERAVRSPENPMDVATRAQGSSTSDVETVMSVADMTKEASVKTAAAV
ATGLISLPSQKTFLKSSSFQKQKQMSMESDVATIGSVRADDSALPRMDATRENI
VSTWQIKSLAPDRIEMLPVLDGRMLKITDRALFAQLGLVDYTLTKLSVDS
VTVSAGQTRALVBAITLESACLSDLVHPENNATDVRYTTRTREVFMKSGMKITTEGSV
LAS"
2407..2436
/gene="At5g42480"
misc_difference 2433.
/gene="At5g42480"
/note="compared to genomic sequence"
/replace="tlg"
ORIGIN
Query Match 99.7%; Score 2399.6; DB 15; Length 2436;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2402; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGGAAGCTGAGTCAAGTCGCGATTGATGTCCTCCCATTCGAATTATGCCATTACCA 60
Db 1 ATGGAAGCTGAGTCAAGTCGCGATTGATGTCCTCCCATTCGAATTATGCCATTACCA 60
QY 61 CCGGCGACGACAAAGCTCCGAGTGAACCAACCTCTTACACTATCTGCTCCGACAGC 120
Db 61 CCGGCGACGACAAAGCTCCGAGTGAACCAACCTCTTACACTATCTGCTCCGACAGC 120
QY 121 AAATGGGCGGACGGTCTCTCTCCGACTTCAATTCACTCCGATTCCTCTCTCTCC 180
Db 121 AAATGGGCGGACGGTCTCTCTCCGACTTCAATTCACTCCGATTCCTCTCTCTCC 180
QY 181 TTGCGCACCGGACCAACCAACCGCACTCTGCTCTCTGCGCACCATCTATGTAGTCGCC 240
Db 181 TTGCGCACCGGACCAACCAACCGCACTCTGCTCTCTGCGCACCATCTATGTAGTCGCC 240
QY 241 GAACGCCACGTCGCCATTCCTCATGATTCTTCAAGGATTTAGAGCTCAAAACATTTTC 300
Db 241 GAACGCCACGTCGCCATTCCTCATGATTCTTCAAGGATTTAGAGCTCAAAACATTTTC 300
QY 301 TTAAACCATGGAATCAGAAAGAGCATTCGAAAGTAGGGTTTCGAAACCGCGCAATTGGT 360
Db 301 TTAAACCATGGAATCAGAAAGAGCATTCGAAAGTAGGGTTTCGAAACCGCGCAATTGGT 360
QY 361 TTGACGACGACGCTTTATATGACCGAGACAGATTTCTTCAAGCTGCTTGGAATCTGTG 420
Db 361 TTGACGACGACGCTTTATATGACCGAGACAGATTTCTTCAAGCTGCTTGGAATCTGTG 420
QY 421 TCTAATCTCGTGTCTAAGAAAGAGTACATGAAGGTCTTCTTGATGTAAAGAAAGTACA 480
Db 421 TCTAATCTCGTGTCTAAGAAAGAGTACATGAAGGTCTTCTTGATGTAAAGAAAGTACA 480
QY 481 GTCAATCTGATGTCCTTGCGGATTAAGGTCCTGGGCTCTCTGATTTGCAAGAAAGGT 540
Db 481 GTCAATCTGATGTCCTTGCGGATTAAGGTCCTGGGCTCTCTGATTTGCAAGAAAGGT 540
QY 541 GGTGAGACTGAGATAGTTCCTCGGGTGTGAGGCTCTGCTTMAAGAGAGTTCCCTAG 600
Db 541 GGTGAGACTGAGATAGTTCCTCGGGTGTGAGGCTCTGCTTMAAGAGAGTTCCCTAG 600
QY 601 TCGTTTAAAGCAAGATGTGTTTATGTAAGGCGCTTGCTTCTGATGTCTGAGGAT 660
Db 601 TCGTTTAAAGCAAGATGTGTTTATGTAAGGCGCTTGCTTCTGATGTCTGAGGAT 660
QY 661 GCTATGGATTTGATTCACCTGATTTTATTAATGCTGTTATGAGTTTGTGAAGAAAGCTTG 720
Db 661 GCTATGGATTTGATTCACCTGATTTTATTAATGCTGTTATGAGTTTGTGAAGAAAGCTTG 720
QY 721 AAGCTTTTAAAGAGAGAAAGAGCAAGTAGCCTTGACCGGATTTTACGTAACAATTGAT 780
Db 721 AAGCTTTTAAAGAGAGAAAGAGCAAGTAGCCTTGACCGGATTTTACGTAACAATTGAT 780
QY 781 GAGACTTTGGAAGAGATCACTCCGCTGATGTCTTGGAGCTACTTGGCTTACCGCTTGGT 840
Db 781 GAGACTTTGGAAGAGATCACTCCGCTGATGTCTTGGAGCTACTTGGCTTACCGCTTGGT 840
QY 841 GATGATTAACGCTGGGAAAGAGCTAAATGGTTTAAAGCGGTGCGGAATATTTTGGGCT 900
Db 841 GATGATTAACGCTGGGAAAGAGCTAAATGGTTTAAAGCGGTGCGGAATATTTTGGGCT 900
QY 901 GTTGGAGAGGTGAGCATCTGTTGTTGGGGTTTGAACCGGTGAGAAATTTATGAT 960
Db 901 GTTGGAGAGGTGAGCATCTGTTGTTGGGGTTTGAACCGGTGAGAAATTTATGAT 960
QY 961 GAGGCGTTTTTTCAGATGACAGCTGCTGAGCAGGTTGATCTTTTGTAGTACCCCAAGC 1020
Db 961 GAGGCGTTTTTTCAGATGACAGCTGCTGAGCAGGTTGATCTTTTGTAGTACCCCAAGC 1020
QY 1021 AATTTTCCAGAGAGTCAATTGAGTTTAAAGTGTGACATTTGCTTGTGGCTCAAGCT 1080
Db 1021 AATTTTCCAGAGAGTCAATTGAGTTTAAAGTGTGACATTTGCTTGTGGCTCAAGCT 1080

QY 1081 TTTATTTGTAAGAGCCACACCTTTTACAGAGTGTGATTAAGCAATTCAGCAACTTCAG 1140
Db 1081 TTTATTTGTAAGAGCCACACCTTTTACAGAGTGTGATTAAGCAATTCAGCAACTTCAG 1140
QY 1141 CAGGCTTAAGGTAATGGCTATGAGGATTTCTGCGCATTTGTATGATACAGGAAATAATGG 1200
Db 1141 CAGGCTTAAGGTAATGGCTATGAGGATTTCTGCGCATTTGTATGATACAGGAAATAATGG 1200
QY 1201 GAGATAGACTTCGGTCTAAGAAAGGGAGCTCTGTCACCTGCTTATAGGCAAAAGTTGATGA 1260
Db 1201 GAGATAGACTTCGGTCTAAGAAAGGGAGCTCTGTCACCTGCTTATAGGCAAAAGTTGATGA 1260
QY 1261 TGCCTATGTGTGTGGGCTTTAGACAGTGAAGATTCAAAATATAGAAATCCAGTATTTGTG 1320
Db 1261 TGCCTATGTGTGTGGGCTTTAGACAGTGAAGATTCAAAATATAGAAATCCAGTATTTGTG 1320
QY 1321 GAGTTTGTTTTGAAGAAATTCGAATTCGTATGACAAAGATGATCTCCCTGACATATGCAAA 1380
Db 1321 GAGTTTGTTTTGAAGAAATTCGAATTCGTATGACAAAGATGATCTCCCTGACATATGCAAA 1380
QY 1381 TTGTTGAAAACTGTGTTGCGAGGGGTTGTCTTCTTCTAGGTTCAAGACACAAAGATTA 1440
Db 1381 TTGTTGAAAACTGTGTTGCGAGGGGTTGTCTTCTTCTAGGTTCAAGACACAAAGATTA 1440
QY 1441 AAATTTAACTCGGGGACTATATGATGATCTTATGATGATCTTATGATTTGAGTTACTTGA 1500
Db 1441 AAATTTAACTCGGGGACTATATGATGATCTTATGATGATCTTATGATTTGAGTTACTTGA 1500
QY 1501 GAGGTAGTTCAAGGTTCTCTTTAGCTGCTGCTGCAATATGGCAAGATTTGAGCCGAG 1560
Db 1501 GAGGTAGTTCAAGGTTCTCTTTAGCTGCTGCTGCAATATGGCAAGATTTGAGCCGAG 1560
QY 1561 CATGTGAAGCTATGCTATGACGAGCACTGACAAAGTTTTCCTCCGCTATACAGAT 1620
Db 1561 CATGTGAAGCTATGCTATGACGAGCACTGACAAAGTTTTCCTCCGCTATACAGAT 1620
QY 1621 AGAACTCGGCTGAAACCAAGATGTGCAAGACAGTGTATGATGATCTCTGTTGGT 1680
Db 1621 AGAACTCGGCTGAAACCAAGATGTGCAAGACAGTGTATGATGATCTCTGTTGGT 1680
QY 1681 AACAATGTAGGCGGTATGAGTGTGAGCTGCTGCTTATGATGACAGAACTGTAAAGCCCTCT 1740
Db 1681 AACAATGTAGGCGGTATGAGTGTGAGCTGCTGCTTATGATGACAGAACTGTAAAGCCCTCT 1740
QY 1741 GAAAACTTTGAAAATATGATTTATGCAATTCGACTGGGCTCTCAGAGAGTACGTTGAT 1800
Db 1741 GAAAACTTTGAAAATATGATTTATGCAATTCGACTGGGCTCTCAGAGAGTACGTTGAT 1800
QY 1801 GAAACTATCTGTGAAAATGTCCGTTGCTGATATGTTAAAGAGGCAAGTGTGAATCTTA 1860
Db 1801 GAAACTATCTGTGAAAATGTCCGTTGCTGATATGTTAAAGAGGCAAGTGTGAATCTTA 1860
QY 1861 GCTGCTGTGTGGCAATTTGACATGATTTTCACTGTTCAGCCAGAAAGTATTTCTTAAAGC 1920
Db 1861 GCTGCTGTGTGGCAATTTGACATGATTTTCACTGTTCAGCCAGAAAGTATTTCTTAAAGC 1920
QY 1921 AGCTCATCTTTTCAAGCAAGAGATATGTTTCTTATGAGAACTGATGTCCGTACATA 1980
Db 1921 AGCTCATCTTTTCAAGCAAGAGATATGTTTCTTATGAGAACTGATGTCCGTACATA 1980
QY 1981 GGGTCAGTCAGACTGACATTCAGAAACCTTCCAGAAATGATCTGAGACTGACAGAG 2040
Db 1981 GGGTCAGTCAGACTGACATTCAGAAACCTTCCAGAAATGATCTGAGACTGACAGAG 2040
QY 2041 AATATATATCCCAAGTGTGCAAGATTTAAAGTCTCTGCTTTTGGGCTGATTCACCGCAT 2100
Db 2041 AATATATATCCCAAGTGTGCAAGATTTAAAGTCTCTGCTTTTGGGCTGATTCACCGCAT 2100
QY 2101 GAAATGTTACCAAGGTTTGTGATGGCGAAATCTGAAGATTTTGAAGTGTGACAGAGAGCT 2160
Db 2101 GAAATGTTACCAAGGTTTGTGATGGCGAAATCTGAAGATTTTGAAGTGTGACAGAGAGCT 2160
QY 2161 GAAACTGCGGACGCTTGGGTTGTTTATGATTAACACTGTTGAAACTATCTGTTGACAGT 2220

Db 2161 GAAATCGCGACGCTGGGTTGATTATGATATACAGTGTGAAAATATCTTGTACACGT 2220
Oy 2221 GAGACAGTCTGACGAGATGGAACCGGTGCTGCTGGAAGAACACTGTGAGAGAGTGTGT 2280
Db 2221 GTGACAGTCTGACGAGATGGAACCGGTGCTGCTGGAAGAACACTGTGAGAGAGTGTGT 2280
Oy 2281 TGTCTATCTGATTTGGTTCATCCAGAAAACAATGCTACTGATGTCAGAACCTACACACA 2340
Db 2281 TGTCTATCTGATTTGGTTCATCCAGAAAACAATGCTACTGATGTCAGAACCTACACACA 2340
Oy 2341 AGATACGAAGTTTCTGCTGCTCAGTCAGGCTGGAATAATCTGAAAGGCTCTGTTCTTGA 2400
Db 2341 AGATACGAAGTTTCTGCTGCTCAGTCAGGCTGGAATAATCTGAAAGGCTCTGTTCTTGA 2400
Oy 2401 TCATTA 2406
Db 2401 TCATTA 2406

RESULT 3
AY091075
LOCUS
DEFINITION
Accession
Version
Keywords
Source
Organism
Reference
Authors

AY091075 2637 bp mRNA linear PLN 18-SEP-2002
Arabidopsis thaliana unknown protein (At5g42480) mRNA, complete
cds.
AY091075.1 GI:20259550
FLI CDS.
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 2637)

Yamada, K., Banh, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M.,
Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L.,
Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L.,
Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,
Jones, T., Kamitaya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,
Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J.,
Sakurai, T., Satou, M., Seki, M., Shim, P., Southwick, A.,
Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.
Arabidopsis Full Length cDNA Clones
Unpublished
2 (bases 1 to 2637)

Yamada, K., Banh, J., Chan, M.M., Chang, C.H., Chang, E., Dale, J.M.,
Deng, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L.,
Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L.,
Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,
Jones, T., Kamitaya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,
Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J.,
Sakurai, T., Satou, M., Seki, M., Shim, P., Southwick, A.,
Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.
Direct Submission
Submitted (21-MAR-2002) Plant Gene Expression Center, 800 Buchanan
Street, Albany, CA 94710, USA
RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAPL cDNAs (RAPL cDNA: 'RIKEN
Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J.,
Satou, M., Kamitaya, A., Sakurai, T., Carninci, P., Kawai, J.,
Hayashizaki, Y. and Shinozaki, K.

The Salk, Stanford, PGSC (SSP) Consortium members carried out the
sequencing and annotation of the RAPL cDNAs: Yamada, K., Banh, J.,
Chan, M.M., Chang, C.H., Chang, E., Dale, J.M., Deng, J.M.,
Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C.,
Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Bowser, L.,
Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,
Jones, T., Kamitaya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Lam, B.,
Lin, J., Meyers, M.C., Miranda, M., Nguyen, M., Palm, C.J., Shim, P.,
Southwick, A., Davis, R.W., Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PGSC) and Seki, M. (RIKEN GSC) contributed equally to
this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGSC)

contributed equally to this work as PIs.

Annotation is based on the January 2002 version of the Arabidopsis
genome submitted to GenBank.

FEATURES

source

Location/Qualifiers

1..2637

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/chromosome="5"

/clone="RAPL09-76-G11 (R19395)"

/ecotype="Columbia"

/note="This clone is in a modified pBluescript vector
(FLC-1) as a BamHI/XhoI insert."

1..2637

/gene="At5g42480"

1..113

/gene="At5g42480"

114..2519

/gene="At5g42480"

/codon_start=1

/evidence=experimental

/product="unknown protein"

/protein_id="AAM13895.1"

/db_xref="GI:20259551"

/translation="MEALSHVIGIGLSPPOLCLPLPATTKLRSHNTSTTICSAKMD
RLSDPNTSDSSSSSPATATTATTATVATVPSISIRPBRHVIPIDFVIGLAOTHELT
DGTIRAPBAPRSPPQRFSDALISRKQIIQAACETISNRSREYREGLLDDEAT
VITVPMDKVPALCVLOEGGETIIVKVGSLAKERIPKPKDQVIVVMAALFIDVS
RDAAALDPDFITGYEFVEBALKLOEGSALAPLDAQIDETLEETTPRVLELG
LPLDDVAAKRLNGISGRNITLMSVGGGAGALVGLRKFPMNEAFIRMALAEQVDL
FVATPSINPAISFVYEVVALVAQAFIGKKFLLDADKQFOOLOQALVAMEIIPAM
LYDRNMWEIDFGIERGICALLIGKVDCEKMWLGLDSESDQYRNPALIVEFVENSND
DNDLRLCCTLEETWLAGVPPRPDRDKFKTGDYDDPMUTSYLSEVVOGSPIL
AALAAARIGAEHYKASMOQLQVFPSPRITDRNSAEKQVETVFSVDPGNVNGRD
GEPVFLAEAPRSPENFETNDYALRAGVSESSVDTEVMSVADMLKEASVXITLAVG
AIGISIFSQEVRPFKSSSFQKDMVSMESDVATIGVRADDSALPRMDARTLENI
VSKKQKISLAFGSDHRIEMPEVLDGMLKITPDRAETQGLVYDTTLKISVDS
VTVAADQTRALVEATLEBSACLSDLVHPENNATVRYTTRVYFVMSGKRTIGSV
LAS"

2520..2637

/gene="At5g42480"

ORIGIN

3'UTR

Query Match 99.7%; Score 2399.6; DB 15; Length 2637;
Best local similarity 99.8%; Pred. No. 0;
Matches 2402; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Oy 1 ATGGAAGCTCGAAGTCAAGTGGGATGGTCTCCCATTCATTAATGCGAATTACCA 60
Db 114 ATGGAAGCTCGAAGTCAAGTGGGATGGTCTCCCATTCATTAATGCGAATTACCA 173
Oy 61 CCGGCGACGACAAAGCTCCGACGTAAGCCAAACCTTACCACTATCTGCTCGCCAGC 120
Db 174 CCGGCGACGACAAAGCTCCGACGTAAGCCAAACCTTACCACTATCTGCTCGCCAGC 233
Oy 121 AAATGGGCGGACGCTTCTCTCGAGCTTCAATTTACCTCGGATTCCTCTCCCTCC 180
Db 234 AAATGGGCGGACGCTTCTCTCGAGCTTCAATTTACCTCGGATTCCTCTCCCTCC 293
Oy 181 TTCCGACCGGACCAACCAACCGCACTCTCTCTGACCAACATCTATTTGATGTCCTC 240
Db 294 TTCCGACCGGACCAACCAACCGCACTCTCTCTGACCAACATCTATTTGATGTCCTC 353
Oy 241 GAATGCCACGTCCTCCATTCCTCCATTTCTTACCAAGATTAAGAGCTCAACACATTC 300
Db 354 GAATGCCACGTCCTCCATTCCTCCATTTCTTACCAAGATTAAGAGCTCAACACATTC 413
Oy 301 TTAACCGATGAATCAGAAAGAGCTTGAAGCTAGGGTTTGAACCGCGCAATTTGGGT 360
Db 414 TTAACCGATGAATCAGAAAGAGCTTGAAGCTAGGGTTTGAACCGCGCAATTTGGGT 473
Oy 361 TTACGCGACGACGCTTATTCAGCCGAGACAGATTTCTTCAAGCTGTCTTGAAACTCTG 420

Db	474	TTCCAGCGACGACGCTTTAATCAGCCGGAGACAGATTCTTCAAGCTGCTTGGAAACTTGG	533
Qy	421	TCTAATCCTCGGCTCAGAAAGAGATACATGAGGCTCTTGTGATGATGAAGAACTTCA	480
Db	534	TCTAATCCTCGGCTCAGAAAGAGATACATGAGGCTCTTGTGATGATGAAGAACTTCA	593
Qy	481	GTCATCATCTGATGTTCCCTGGATTAAGTTCTCTGGGCGCTCTCTGTGTATTTGCAGAGGT	540
Db	594	GTCACTACCTGAAAGTTCCTTGGGATTAAGGTTCTGTGTGCTCTCTGTGTATTTCAAGAAAGT	653
Qy	541	GGTGAAGCTGAGATGATGTTCTTCCGGTGGTGAAGGCTGCTTAAGGAAGGTTGGCTAAG	600
Db	654	GGTGAAGCTGAGATGATGTTCTTCCGGTGGTGAAGGCTGCTTAAGGAAGGTTGGCTAAG	713
Qy	601	TGCTTTAAGCAGATGATGTTTAACTTATATGCGCTTGCGCTTCTCATGTCTCGAGGAT	660
Db	714	TGCTTTAAGCAGATGATGTTTAACTTATATGCGCTTGCGCTTCTCATGTCTCGAGGAT	773
Qy	661	GCTATGCAATGAGATCCACCTGATTTTATTACTGGTATAGATTGTTGAAGAAAGCTTGG	720
Db	774	GCTATGCAATGAGATCCACCTGATTTTAACTGGTATAGATTGTTGAAGAAAGCTTGG	833
Qy	721	AAAGCTTTTAAACGAGAGAAAGAGCAATGAGCTTGAACCGGATTAACGTGACCAATTTGAT	780
Db	834	AAAGCTTTTAAACGAGAGAAAGAGCAATGAGCTTGAACCGGATTAACGTGACCAATTTGAT	893
Qy	781	GAGACTTTGGAAGAGATCATCTCCGCTTATGCTTGAAGACTTACTTGCTTACCGCTTGGT	840
Db	894	GAGACTTTGGAAGAGATCATCTCCGCTTATGCTTGAAGACTTACTTGCTTACCGCTTGGT	953
Qy	841	GATGATTACGCTGCGAATAAGATTAATGATTAAGCGGCTGTCGGAATATTTTGTGCTCT	900
Db	954	GATGATTACGCTGCGAATAAGATTAATGATTAAGCGGCTGTCGGAATATTTTGTGCTCT	1013
Qy	901	GTTGAGAGAGGTGGAGCATACAGCTCTTGTGGGGTTTGAACCCGGAAGATTTATAGAT	960
Db	1014	GTTGAGAGAGGTGGAGCATACAGCTCTTGTGGGGTTTGAACCCGGAAGATTTATAGAT	1073
Qy	961	GAGGCGTTTAAACGATACAGCTGCTGAGCAGGTTGATCTTTTGTGTACACCCAAAGC	1020
Db	1074	GAGGCGTTTAAACGATACAGCTGCTGAGCAGGTTGATCTTTTGTGTACACCCAAAGC	1133
Qy	1021	AATATTCCAGACAGATGATTTTGAATTTACGAATTTGACATTTGCTCTTGTGGCTCAAGCT	1080
Db	1134	AATATTCCAGACAGATGATTTTGAATTTACGAATTTGACATTTGCTCTTGTGGCTCAAGCT	1193
Qy	1081	TTTATTTGTAAGAAACCAACCTTTTACAGATGCTGATTAAGCAATTCACACCACTTCAG	1140
Db	1194	TTTATTTGTAAGAAACCAACCTTTTACAGATGCTGATTAAGCAATTCACACCACTTCAG	1253
Qy	1141	CAGGCTTAAGTAAATGGCTATGAGAGATTCCTGCGATGTTGTATGATACACGAAATTAATGG	1200
Db	1254	CAGGCTTAAGTAAATGGCTATGAGAGATTCCTGCGATGTTGTATGATACACGAAATTAATGG	1313
Qy	1201	GAGATAGAATTCGAGTCTAGAAAGGGGACTCTGTGCACTCTTATTAAGCAAGTTGATGA	1260
Db	1314	GAGATAGAATTCGAGTCTAGAAAGGGGACTCTGTGCACTCTTATTAAGCAAGTTGATGA	1373
Qy	1261	TGCCGTATGTGTTGGGCTTACAGACGTGAGATTACAAATTAAGAAATCCAGTATTTGTG	1320
Db	1374	TGCCGTATGTGTTGGGCTTACAGACGTGAGATTACAAATTAAGAAATCCAGTATTTGTG	1433
Qy	1321	GAGTTTGTTTGGAAATTCAAATGTGTATGACATGATGATCTCCCTGGACTATGCAAA	1380
Db	1434	GAGTTTGTTTGGAAATTCAAATGTGTATGACATGATGATCTCCCTGGACTATGCAAA	1493
Qy	1381	TTGTGTGAAACCTGTGTGGACAGGGGTTGTCTTCTTAGGTTCAAGACACCAAAAGATTA	1440
Db	1494	TTGTGTGAAACCTGTGTGGACAGGGGTTGTCTTCTTAGGTTCAAGACACCAAAAGATTA	1553
Qy	1441	AAATTTAACTCGGGGACTACTATGATATCTATGTATTTTGAAGTTACTTGAAGAAAGTGT	1500

Db	1554	AAATTAAACTCGGGGACACTACTATGATGATTCCTATGGTTTGAAGTACTTGGAAAGAGCTG	1613
Qy	1501	GAGGTAGTCAAGGGTTCCTCTTATAGCTGCTGCGCACTATGGCGAAGATTGAGCCGAG	1560
Db	1614	GAGGTAGTCAAGGGTTCCTCTTATAGCTGCTGCGCACTATGGCGAAGATTGAGCCGAG	1673
Qy	1561	CATCTGAAAGCTAGTGGTCTATGCAAGGCACTGCAGAAATTTTCTCCCGTATACAGAT	1620
Db	1674	CATCTGAAAGCTAGTGGTCTATGCAAGGCACTGCAGAAATTTTCTCCCGTATACAGAT	1733
Qy	1621	AGAAACTCGGCTGAACCCAGATGTCGAAGACAGCTGTTAGTGAATCCTGTTGAT	1680
Db	1734	AGAAACTCGGCTGAACCCAGATGTCGAAGACAGCTGTTAGTGAATCCTGTTGAT	1793
Qy	1681	AACAAATGAGCCGTGATAGTGTGAGCTTGCTCTTATTTGCAAGAGCTGTAGACCTCT	1740
Db	1794	AACAAATGAGCCGTGATAGTGTGAGCTTGCTCTTATTTGCAAGAGCTGTAGACCTCT	1853
Qy	1741	GAAACCTTGAACCTAATATGATTTATGCAATTCGAGTCGGGGTCTCAGAGATGAGCTGAT	1800
Db	1854	GAAACCTTGAACCTAATATGATTTATGCAATTCGAGTCGGGGTCTCAGAGATGAGCTGAT	1913
Qy	1801	GAAACTACTGTGAAATATGTCGGTGTCTGATATGTTAAAGGAGCGAAGTGTGAATCCTA	1860
Db	1914	GAAACTACTGTGAAATATGTCGGTGTCTGATATGTTAAAGGAGCGAAGTGTGAATCCTA	1973
Qy	1861	GCTGCTGGTGTGCGCAATTGGACTGATATTTCACTGTTCACGCCAAGATATTTTCTTAAAGC	1920
Db	1974	GCTGCTGGTGTGCGCAATTGGACTGATATTTCACTGTTCACGCCAAGATATTTTCTTAAAGC	2033
Qy	1921	AGCTCATCTTTTCAAGCGCAAGGATATGATGTTCTTATGGAATCTGATGTGCGTACCATTA	1980
Db	2034	AGCTCATCTTTTCAAGCGCAAGGATATGATGTTCTTATGGAATCTGATGTGCGTACCATTA	2093
Qy	1981	GGGTCACTGACAGCTGACGATTCAGAAAGCATTTCCAGAAATGGAATGCTAGAGACTGCAGAG	2040
Db	2094	GGGTCACTGACAGCTGACGATTCAGAAAGCATTTCCAGAAATGGAATGCTAGAGACTGCAGAG	2153
Qy	2041	AATATAGTATTCGAAGTGGCAGAAAGATTAACTCTCTGCTTTTGGGCGCTGATCACCCGATA	2100
Db	2154	AATATAGTATTCGAAGTGGCAGAAAGATTAACTCTCTGCTTTTGGGCGCTGATCACCCGATA	2213
Qy	2101	GAAATGTTAACAGAGTTTTGGATGGGCGAATGCTGAAGATTTGGACTGACAGAGCAGCT	2160
Db	2214	GAAATGTTAACAGAGTTTTGGATGGGCGAATGCTGAAGATTTGGACTGACAGAGCAGCT	2273
Qy	2161	GAAACTGCGCAGCTTGGGTTGTTATGATTTATACACTGTGTAACATATCTGTGAACAGT	2220
Db	2274	GAAACTGCGCAGCTTGGGTTGTTATGATTTATACACTGTGTAACATATCTGTGAACAGT	2333
Qy	2221	GTCGCACTTCCAGAGATGGAACCCCGTCTCTGTGGAACAACCTCTGGAGAGTCTGCT	2280
Db	2334	GTCGCACTTCCAGAGATGGAACCCCGTCTCTGTGGAACAACCTCTGGAGAGTCTGCT	2393
Qy	2281	TGTTATCTGATTTGGTTCATCCAGAAAACATATGCTACTATGTCABAACCTACACACA	2340
Db	2394	TGTTATCTGATTTGGTTCATCCAGAAAACATATGCTACTATGTCABAACCTACACACA	2453
Qy	2341	AGATAGCAAGTTTTCTGCTCCAACTCAGGCTGAGAAAATCACTGAAGGCTCTGTTCTTGA	2400
Db	2454	AGATAGCAAGTTTTCTGCTCCAACTCAGGCTGAGAAAATCACTGAAGGCTCTGTTCTTGA	2513
Qy	2401	TCATATA 2406	
Db	2514	TCATATA 2519	

RESULT 4					
AY221468					
LOCUS	AY221468	3668 bp	DNA	linear	PLN 05-AUG-2003
DEFINITION	Arabidopsis thaliana division protein (ARC6) gene, complete cds;				
	nuclear gene for chloroplast product.				
ACCESSION	AY221468				

VERSION	AY221468.1	GI:33436338
KEYWORDS		
SOURCE	Arabidopsis thaliana (thale cress)	
ORGANISM	Arabidopsis thaliana	
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Euparietia; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.	
REFERENCE	1 (bases 1 to 3668)	
AUTHORS	Vitha,S., Froehlich,J.E., Koksharova,O., Pyke,K.A., Van Ert,H. and OsterYoung,K.W.	
TITLE	ARC6 Is a J-Domain Plastid Division Protein and an Evolutionary Descendant of the Cyanobacterial Cell Division Protein Ftn2	
JOURNAL	Plant Cell 15 (8), 1918-1933 (2003)	
PUBMED	12897262	
REFERENCE	2 (bases 1 to 3668)	
AUTHORS	Vitha,S., Koksharova,O., van Ert,H., Froehlich,J.E. and OsterYoung,K.W.	
TITLE	Direct Submission	
JOURNAL	Submitted (17-JAN-2003) Department of Plant Biology, Michigan State University, 166 Plant Biology Building, East Lansing, MI 48824, USA	
FEATURES	Location/Qualifiers	
SOURCE	1..3668	

```

mRNA
join(<478, .984,1067,.1291,1379. .1639,1723. .2710,
3000. .3133. 3225. .>3515)
/gene="ARC6"
/product="division protein"
join(478. .984,1067. .1291,1379. .1639,1723. .2710,3000. .3133
join(478. .984,1067. .1291,1379. .1639,1723. .2710,3000. .3133
CDS

```

```

/gene="AKC6"
/note="inner envelope membrane-localized; similar to
synchobacterial cell division protein sli10169 of
Synchococcus; contains J-domain"
/codon_start=1
/product="division protein"
/protein_id="AAQ18645.1"
/db_xref="GI:33436339"
/translation="MEALSHVGIQLSPQCRLLPAPPTKLRSHNSTTTI CSAKND
RLSDPFTSDSSSSPATAITTAITLTVSLPSIDIRPHNPIDVQVQATPHLT
DGRRAFEARVSKPQFGESDDALISRQILLQACETLSNPSRREYNBGLDEEAT
VITDVPDKVPGALCVLQEGEGETEIVLRVEALLKEKLSPSFQDVLVWALAFLDVS
RDMALDPPDPIITGFEVEBALKILQEGGASLAVGLDRAQIDITLIEITPRVYIELLG
LPIEDVYARKLINGLISGVRIIMWSVGGASLAVGLDRAQIDITLIEITPRVYIELLG
FVATPSNI PASFEVVEVALVAQAFITGKPHILQDADQPOQLAQKMAEIPAM
LYTRRNWEIDPGLERGLCALLIGKDECHMILQBSDESYRNALIVEVLENSRD
DNDLD,PELCLLETWLAGVVFPRFRDTRKFFKGLGDDYDDPMWLSIYERVEVQGSPL
AAATATARI GAEHYKASAMQALQKVPESRTIDRNSAPKQFETVDSVGNVNGVD
GEEGVFIAELAVRESENDETIDYALRAGVSSYDDEITVEGVSMDLKEASVQILLAGV
AIGLISLFQKYLFKSSSFQKQDOWMSMSDVAITGVARADSBALPRDATTANI
VSMQWIKSLAPGPDHRIEMLPVLQDGRMKIKITWDYRAETRAQGLVYDITLGLSDS
VTVSAGSTRALVEATLIESACLSDLVHPENNATIDVRYITTYREVEFWKSGMKITGESV
LAS"

```

Query Match	71.7%	Score 1724	DB 15	Length 3668
Best Local Similarity	79.2%	Pred. No. 0		
Matches 2406	Conservative 0	Mismatches 0	Indels 632	Gaps 5
QY	1	ATGGAAGCTCTGAGTCAAGTCGCGCATGGTGTCTCCCCCATTCGAAATATGCGGATTAACA	60	
Db	478	ATGGAAGCTCTGAGTCAAGTCGCGCATGGTGTCTCCCCCATTCGAAATATGCGGATTAACA	537	
QY	61	CCGGCGACGACAAAGCTCCGACGTAGCCACAACCTCTACAACTATCTCTCCGCGAGC	120	
Db	538	CCGGCGACGACAAAGCTCCGACGTAGCCACAACCTCTACAACTATCTCTCCGCGAGC	597	
QY	121	AAATGGCGCGACCGCTCTCTCTCCGACCTTCAATTTCACTCCGATTTCTCTCTCTCTCC	180	

D	598	AAATGGGCGACCGCTTCTCTCGGACTTCAATTCACTCGGATTTCCCTCCTCCTCC	657
O	181	TTCCGCAACCGCACCAACACCGCACTCTCGTCTCTGCGCAACATCTATTGATGTGCC	240
D	658	TTCCGCAACCGCACCAACACCGCACTCTCGTCTCTGCGCAACATCTATTGATGTGCC	717
O	241	GAAAGCCACGCTCCCATCTCCCATGATTCTTACACGAGTATTAGAGCTCAACACATTTTC	300
D	718	GAAAGCCACGCTCCCATCTCCCATGATTCTTACACGAGTATTAGAGCTCAACACATTTTC	777
O	301	TTAAACGATGGAATCAGAAAGACTTCGACCTAGGCTTTCGAAACCGCCGCAATTCCGT	360
D	778	TTAAACGATGGAATCAGAAAGACTTCGACCTAGGCTTTCGAAACCGCCGCAATTCCGT	837
O	361	TTTCAGCCACGCTTTTAATCAGCCCGGAGACAGATTCTTCAAGCTGCTTGGAAACTCTG	420
D	838	TTTCAGCCACGCTTTTAATCAGCCCGGAGACAGATTCTTCAAGCTGCTTGGAAACTCTG	897
O	421	TCCTATCTCCGCTTAGAAGAGACTCAATGAAGTCTTCTTGATGATGAAGAAGCTACA	480
D	898	TCCTATCTCCGCTTAGAAGAGACTCAATGAAGTCTTCTTGATGATGAAGAAGCTACA	957
O	481	GTCATCATCTGATGTTCTTGGGAT-----	504
D	958	GTCATCATCTGATGTTCTTGGGATTAAGGTAATTCGATTCGATAATAAATCTTTC	1011
O	505	-----AAGTTCCTGGGGC	518
D	1018	GTTTAAATTTCATGAATTGATTAAGAGAGACTTTTATCTAGTGAAGGTTCTCGGGGC	1077
O	519	TCTCTGTGTATTGCAAGAGGTGTGTGACTGATAGTCTTCCGGTGTGTGAGGCTCT	578
D	1078	TCTCTGTGTATTGCAAGAGGTGTGTGACTGATAGTCTTCCGGTGTGTGAGGCTCT	1133
O	579	GCTTAAGAGAGTGTGCTTAAGTGTGTTAAAGACAGATGTGTTTTAGTTATGGCGCTTGC	638
D	1138	GCTTAAGAGAGTGTGCTTAAGTGTGTTAAAGACAGATGTGTTTTAGTTATGGCGCTTGC	1197
O	639	GTTTCTGATGTCTCAGAGGATGTCTATGGCACTTGGATCCACTGATTTTATTAACGTGTA	698
D	1198	GTTTCTGATGTCTCAGAGGATGTCTATGGCACTTGGATCCACTGATTTTATTAACGTGTA	1255
O	699	TGAGTTTGTGAGAGACTTTGAAGCTTTTAC-----	730
D	1258	TGAGTTTGTGAGAGACTTTGAAGCTTTTACAGGTAGTTTGACTTGCTTTGTAAATTTG	1311
O	731	-----A	731
D	1318	ACGAGCCTTGACTTATTAAGAACTTCTTGATTTGATTACTTGTATTTAGCTCTGTGTGA	1377
O	732	GGAAGAAAGACAGTAGTAGCTTGTGACCGGATTTACGTGACAATAATGATGAGACTTTGGA	791
D	1378	GGAAGAAAGACAGTAGTAGCTTGTGACCGGATTTACGTGACAATAATGATGAGACTTTGGA	1433
O	792	AGAGATCACTCCGCTTATGTCTTGAAGCTAATTGGCTTACCGCTTGGTGTGATTAAGC	851
D	1438	AGAGATCACTCCGCTTATGTCTTGAAGCTAATTGGCTTACCGCTTGGTGTGATTAAGC	1497
O	852	TGCGAAAAGACTAAATGGTTTAAAGCGGTGTGCGGAATATTTTGTGTGTGAGGAGAG	911
D	1498	TGCGAAAAGACTAAATGGTTTAAAGCGGTGTGCGGAATATTTTGTGTGTGAGGAGAG	1557
O	912	TGAGCATCAGCTCTTGTGTGGGGTTTGAACCCGTGAGAACTTATGATGAGGCGTTTT	971
D	1558	TGAGCATCAGCTCTTGTGTGGGGTTTGAACCCGTGAGAACTTATGATGAGGCGTTTT	1617
O	972	ACGAATGACAGCTGTGAGC-----	991
D	1618	ACGAATGACAGCTGTGAGGAGGTATACAGTTTAGATCTTTTAAATTTCTTTAGCA	1677
O	992	-----AGGTGATCTTTTGTGA	1008

Db	1678	TGATATAACTTTAGATTTCATATTTTAATGATGTTGTGTGTAGTTGATCTTTTGTGA	1737
Qy	1009	GCTACCCCAAGCAATATTCACAGAGAATCATTTGAAGTTTACGAAGTTGCACTTGCTCTT	1068
Db	1738	GCTACCCCAAGCAATATTCACAGAGAATCATTTGAAGTTTACGAAGTTTGCACTTGCTCTT	1797
Qy	1069	GTGGCTCAAGCTTTTATTTTGTGAAGAAGCAACCTTTTACAGATGCTGATTAAGCAATTC	1128
Db	1798	GTGGCTCAAGCTTTTATTTTGTGAAGAAGCAACCTTTTACAGATGCTGATTAAGCAATTC	1857
Qy	1129	CAGAACCTTCACAGAGCTTAAGGTATGGCTATGAGATTCCTGGCAGTGTGATATATACA	1188
Db	1858	CAGAACCTTCACAGAGCTTAAGGTATGGCTATGAGATTCCTGGCAGTGTGATATATACA	1917
Qy	1189	CGAATAATTTGGAGATATAGACTTCGCTGTAGAAAAGGGACCTCTGTGCACCTGTATAGGC	1248
Db	1918	CGAATAATTTGGAGATATAGACTTCGCTGTAGAAAAGGGACCTCTGTGCACCTGTATAGGC	1977
Qy	1249	AAAGTTGATGAAATCCCTATGTTGGTGTGGGCTTAACACATGAGATTCACAAATATAGGAT	1308
Db	1978	AAAGTTGATGAAATCCCTATGTTGGTGTGGGCTTAACACATGAGATTCACAAATATAGGAT	2037
Qy	1309	CCAGCTATTGGAGAGTTGTTTGGAGAAATCAATCGTATGCAAAATGATGATCTCCCT	1368
Db	2038	CCAGCTATTGGAGAGTTGTTTGGAGAAATCAATCGTATGCAAAATGATGATCTCCCT	2097
Qy	1369	GGACTATGCAAAATTTGTGGAAACCTGTGTTGCAGAGGGTTGTCTTTCCTAGGTTTCAGAGAC	1428
Db	2098	GGACTATGCAAAATTTGTGGAAACCTGTGTTGCAGAGGGTTGTCTTTCCTAGGTTTCAGAGAC	2157
Qy	1429	ACCAAAAGTAAAAAATTTTAACTCGGGGACATCATATGATTCCTATGCTTTTGAATTAC	1488
Db	2158	ACCAAAAGTAAAAAATTTTAACTCGGGGACATCATATGATTCCTATGCTTTTGAATTAC	2217
Qy	1489	TTGGAAAAGTGGAGGTAGTTCAGGGTCTCTCTTATGCTGCTGTCGACAACTATGCGAAGG	1548
Db	2218	TTGGAAAAGTGGAGGTAGTTCAGGGTCTCTCTTATGCTGCTGTCGACAACTATGCGAAGG	2277
Qy	1549	ATTGGAGCCGAGCATGTGAAAAGCTATGECTATGCAAGGCACTGCAGAAAAGTTTTCCTTCC	1608
Db	2278	ATTGGAGCCGAGCATGTGAAAAGCTATGECTATGCAAGGCACTGCAGAAAAGTTTTCCTTCC	2337
Qy	1609	CGCTATACAGATPAGAACTCGGCTGAAACCAAGATGTGCAGAAGACAGTGTTAAGTGA	1668
Db	2338	CGCTATACAGATPAGAACTCGGCTGAAACCAAGATGTGCAGAAGACAGTGTTAAGTGA	2397
Qy	1669	GATCCTGTGTGTAAACAATGATAGGCCGCGATAGCTGCTGCTTATTTGCGAAGACT	1728
Db	2398	GATCCTGTGTGTAAACAATGATAGGCCGCGATAGCTGCTGCTTATTTGCGAAGACT	2457
Qy	1729	GTAAGACCCCTGTGAAAACCTTGGAAACTATGATTAATGCAATTCAGCTGGGGTCTCAGAG	1788
Db	2458	GTAAGACCCCTGTGAAAACCTTGGAAACTATGATTAATGCAATTCAGCTGGGGTCTCAGAG	2517
Qy	1789	AGTAGCGTTGATGAAACTACTCTGTGAAATGTCCGTTGCTGATATGTTAAAGAGGCAAGT	1848
Db	2518	AGTAGCGTTGATGAAACTACTCTGTGAAATGTCCGTTGCTGATATGTTAAAGAGGCAAGT	2577
Qy	1849	GTTGAAGATCCCTAGCTGCTGTGTGTGCGAATTTGACATTTTACATGTTCCAGCBAABAATAT	1908
Db	2578	GTTGAAGATCCCTAGCTGCTGTGTGTGCGAATTTGACATTTTACATGTTCCAGCBAABAATAT	2637
Qy	1909	TTTCTTAAAGACAGCTCATCTTTTCAACGCAAGATATGTTTTCTTCTATGGAATCTGAT	1968
Db	2638	TTTCTTAAAGACAGCTCATCTTTTCAACGCAAGATATGTTTTCTTCTATGGAATCTGAT	2697
Qy	1969	GTGCTACCA-----1978	
Db	2698	GTGCTACCAATAGGATGATTAATAATGATGCAATTTTCAATATCTGCATTCGCAAAATA	2757
Qy	1979	-----1978	
Db	2758	TGCTGTGTTTGTGAGCTAAGAACATAGTTTCCCACTTATATCATGTCGCCAAAAGTTGTACC	2817

QY	1979	-----	1978
Db	2818	AMGATTAACAAGTTGCTGATGAATTAATTCACATAATATGCTGTAATTTTGTGATCAAA	28777
QY	1979	-----	1978
Db	2878	CTGTAGACAGAAATGTAAATTTCACTCTCAACATTTCTGTTTAGAATACAGGATTAAG	2937
QY	1979	-----	1979
Db	2938	AGATTGCCCTTAGTGTGGCTTGTGCCAATTTTCTTCCCTTGATTTTTCCTTTTCGATTT	2997
QY	1980	AGGTCAGCTGACGACCTACGATTTAGAGAGACCTCCAGAAATGGATACCTAGATGACAGA	2039
Db	2998	AGGGTCAGTCAAGACCTACGATTTAGAGAGACCTTCCAGAAATGGATCTAGAGATGACAGA	3057
QY	2040	GAATATAGTATCCCAAGTGGCAGAGATTAAGTCTGTGGCTTTTGGGCTGTGACCCGCAT	2099
Db	3058	GAATATAGTATCCCAAGTGGCAGAGATTAAGTCTGTGGCTTTTGGGCTGTGACCCGCAT	3117
QY	2100	AGAAATGTTACGAG-----	2113
Db	3118	AGAAATGTTACGAGAGGTGAGGAAATTAATCTACAAATTCATCAATTTGTGAAAACGTG	3177
QY	2114	-----	AGGTTTGGATGGG
Db	3178	TGCACTGATTAATTAAGTCTGCTGCTTGTGATTCTGTATTATTAATTAAGGTTTGGATGGG	3237
QY	2129	GAATGCTGAAGATTGTGACCTGACACAGACAGCTGAACTGGCAGCTTGGGTTGTTATG	2188
Db	3238	GAATGCTGAAGATTGTGACCTGACACAGACAGCTGAACTGGCAGCTTGGGTTGTTATG	3297
QY	2189	ATTATTAACCTGTTGAAAACCTATCTGTTGACAGTGTGACAGTCTCAGCGATGGAAACCCGCTG	2248
Db	3298	ATTATTAACCTGTTGAAAACCTATCTGTTGACAGTGTGACAGTCTCAGCGATGGAAACCCGCTG	3357
QY	2249	CTCTGGTGGAAAGCAACCTGTGGAGAGTGTGCTTCTATCTGATTTGTTGTCATCCAGAAA	2308
Db	3358	CTCTGGTGGAAAGCAACCTGTGGAGAGTGTGCTTCTATCTGATTTGTTGTCATCCAGAAA	3417
QY	2309	ACAATGCTACTGATGTGAGAACTTACACAAACAGATACGAAGTTTGTGTCCAGTCCAG	2368
Db	3418	ACAATGCTACTGATGTGAGAACTTACACAAACAGATACGAAGTTTGTGTCCAGTCCAG	3477
QY	2369	GGTGGAAAATCTACTGAAAGGCTCTGTTCTTGTCATATTA	2406
Db	3478	GGTGGAAAATCTACTGAAAGGCTCTGTTCTTGTCATATTA	3515
RESULT 5			
LOCUS	AY221467	3664 bp	DNA
DEFINITION	Arabidopsis thaliana truncated division protein (arc6) gene, arc6-1 allele, complete cds; nuclear gene for chloroplast product.		
ACCESSION	AY221467		
VERSION	AY221467.1	GI:33436274	
KEYWORDS			
SOURCE	ORGANISM		
	Arabidopsis thaliana (thale cress)		
	Eukaryotes; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.		
REFERENCE	1 (bases 1 to 3664)		
AUTHORS	Vitina,S., Froehlich,J.E., Kokscharova,O., Pyke,K.A., Van Ert,H. and OsterYoung,K.W.		
TITLE	ARC6 is a J-Domain Plastid Division Protein and an Evolutionary Descendant of the Cyanobacterial Cell Division Protein Ftn2		
JOURNAL	Plant Cell 15 (8), 1918-1933 (2003)		
PUBMED	12897262		
REFERENCES	2 (bases 1 to 3664)		
AUTHORS	Vitina,S., Kokscharova,O., van Ert,H., Froehlich,J.E. and OsterYoung,K.W.		

TITLE Direct Submission
JOURNAL Submitted (17-JAN-2003) Department of Plant Biology, Michigan State University, 166 Plant Biology Building, East Lansing, MI 48824, USA
FEATURES
source

1.3664
/organism="Arabidopsis thaliana"

/mol_type="genomic DNA"

/db_xref="taxon:3702"

/ecotype="Ws-2"

/note="genotype: arc6 mutant"

1.3664

/gene="arc6"

/allele="arc6-1"

join(478..984,1067..1291,1379..>1621)

/gene="arc6"

/allele="arc6-1"

/product="truncated division protein"

join(478..984,1067..1291,1379..1621)

/gene="arc6"

/note="truncation caused by premature stop codon"

/codon_start=1

/product="truncated division protein"

/protein_id="AA018644.1"

/db_xref="GI:33436275"

/translation="MEALSHVIGLSPEOLCRLPPATTKLRSHNTSTTIGSASKMA
RLSDPFTSDSSSSPATAATTAATVATLPSIDREHVPIDPYOVLAGOTFLT
DGIRAFARVSKPFGSPSDALISROIIQAACETLSIPRSREVRVGLDDEAT
VITVPMKVRKALCVLOBGSETRIVLVAGELAKERIPKSKQDVIVMLAPLDS
RDMAALDPDFITGIEVEALKEKLQEGASLADPLAQIDETLEITTPRVLELIG
LPUGDYAKRLINGISVRNIIWSVGGSASALVGLTRERKMEAL"

ORIGIN

Query Match 71.5%; Score 1720.8; DB 15; Length 3664;
Best Local Similarity 79.1%; Pred. No. 0;
Matches 2404; Conservative 0; Mismatches 2; Indels 632; Gaps 5;

```

Oy 1 ATGGAAGCTGTGAGTACAGTCGGCATTGGTCTCTCCCAATTCATTATGCGATTACCA 60
Db 478 ATGGAAGCTGTGAGTACAGTCGGCATTGGTCTCTCCCAATTCATTATGCGATTACCA 537
Oy 61 CCGGAGAGCAAAAGCTCGAGCTGACCAACAACCTTCAACATATCTGCTCCGCCAGC 120
Db 538 CCGGAGAGCAAAAGCTCGAGCTGACCAACAACCTTCAACATATCTGCTCCGCCAGC 597
Oy 121 AAATGGCGCAACCGCTTCTCTCGACTTCAATTTCACTCCGATTCCTCCCTCCCTCC 180
Db 598 AAATGGCGCAACCGCTTCTCTCGACTTCAATTTCACTCCGATTCCTCCCTCCCTCC 657
Oy 181 TTGCGCAACCGCAACCGCAACCGCACTCTGCTCTGCGACCACTATTTGATGATGCC 240
Db 658 TTGCGCAACCGCAACCGCAACCGCACTCTGCTCTGCGACCACTATTTGATGATGCC 717
Oy 241 GAAGCGCACTGCCATCCCATTTGATTTCTACAGGATTTAGAGCTTCAACACATTTT 300
Db 718 GAAGCGCACTGCCATCCCATTTGATTTCTACAGGATTTAGAGCTTCAACACATTTT 777
Oy 301 TTAACCAATGAAATGAGAGAGCATTTGAAAGCTAGAGTTTGAACCGCGCAATTCGGT 360
Db 778 TTAACCAATGAAATGAGAGAGCATTTGAAAGCTAGAGTTTGAACCGCGCAATTCGGT 837
Oy 361 TTACGCGAGAGAGCTTTTATATCAGCCGAGAGACAGATTTCTTCAAGCTGTGCGAACTCTG 420
Db 838 TTACGCGAGAGAGCTTTTATATCAGCCGAGAGAGAGATTTCTTCAAGCTGTGCGAACTCTG 897
Oy 421 TCTATCTCTGGTCTAGAGAGAGATGATGAAAGTCTTTGATGATGAAAGAGCTACA 480
Db 898 TCTATCTCTGGTCTAGAGAGAGATGATGAAAGTCTTTGATGATGAAAGAGCTACA 957
Oy 481 GTCATCACTGATGTTCTTTGGGAT----- 504
Db 958 GTCATCACTGATGTTCTTTGGGATAGGTAATTTGCGAATATAAAGTTTCTTC 1017

```

```

Oy 505 -----AAGTTCCTGGGCG 518
Db 1018 GTTTATATTCATGAAATGATTAAGAGAGAACTTTATCTAGAGAGGTTCCCGGGCG 1077
Oy 519 TCTCTGTATTTGCAAGAGAGTGGTGAAGATGATGATTTCTTGGGTTGGTGAAGCTCT 578
Db 1078 TCTCTGTATTTGCAAGAGAGTGGTGAAGATGATGATTTCTTGGGTTGGTGAAGCTCT 1137
Oy 579 GCTTAAGAGAGTGGTGGCTTAAGTGGTTTGAAGAGATGATGATGATGATGATGATGATG 638
Db 1138 GCTTAAGAGAGTGGTGGCTTAAGTGGTTTGAAGAGATGATGATGATGATGATGATGATG 1197
Oy 639 GTTCTCGATGCTTCGAGAGAGTGGTGAAGATGATGATGATGATGATGATGATGATGATG 698
Db 1198 GTTCTCGATGCTTCGAGAGAGTGGTGAAGATGATGATGATGATGATGATGATGATGATG 1257
Oy 699 TGAGTTGTTGAGAGAGCTTTGAAGCTTTTAC----- 730
Db 1258 TGAGTTGTTGAGAGAGCTTTGAAGCTTTTACAGGTAGTTGACTTGGTGTAAATTTG 1317
Oy 731 -----A 731
Db 1318 ACGAGCGTGGCTTTATTAAGAACTTCTGATTTGATTTGATTTGATTTGATTTGATTTG 1377
Oy 732 GAGAGAGAGCAAGTACCTTTCACCGGATTTTACGTCACAAATGATGAGACTTTGGA 791
Db 1378 GAGAGAGAGCAAGTACCTTTCACCGGATTTTACGTCACAAATGATGAGACTTTGGA 1437
Oy 792 AGAATCACTCCGCTTATGCTTTGAGACTTCTGGCTTACCCTGGTGTATGATTAAGC 851
Db 1438 AGAATCACTCCGCTTATGCTTTGAGACTTCTGGCTTACCCTGGTGTATGATTAAGC 1497
Oy 852 TGCGAAAAGCTAATGATTTAGCGGAGTGGCAATTTTGTGCTGTGAGAGAG 911
Db 1498 TGCGAAAAGCTAATGATTTAGCGGAGTGGCAATTTTGTGCTGTGAGAGAG 1557
Oy 912 TGAGCATCAGCTTGTGTGGGGTTTGACCGGTGAGAGTTTATGATGAGGCGTTTT 971
Db 1558 TGAGCATCAGCTTGTGTGGGGTTTGACCGGTGAGAGTTTATGATGAGGCGTTTT 1617
Oy 972 ACGAATGACAGCTGCTGAGC----- 991
Db 1618 ACGAATGACAGCTGCTGAGAGTTATACGTTAGATACCTTTTAAATTTCTTTAGCA 1677
Oy 992 -----AGTTGATCTTTTGTGA 1008
Db 1678 TGATATACTTTAGATTTCTTATTTATGATTTGTTGTTGTTGATTTGATTTGTTGTTG 1737
Oy 1009 GCTACCCCAAGCAATATTCAGAGAGTCAATTTGAGATTTACGAATGTCATTTGCTCTT 1068
Db 1738 GCTACCCCAAGCAATATTCAGAGAGTCAATTTGAGATTTACGAATGTCATTTGCTCTT 1797
Oy 1069 GTGGCTCAAGCTTTTATTTGTTAGAGAGCAACCTTTTACAGATGCTGATTAAGCAATTC 1128
Db 1798 GTGGCTCAAGCTTTTATTTGTTAGAGAGCAACCTTTTACAGATGCTGATTAAGCAATTC 1857
Oy 1129 CAGCAACTTGACAGAGGCTAAGTTATGCTATGAGATTTCTGCGAGTTTGTATGATACA 1188
Db 1858 CAGCAACTTGACAGAGGCTAAGTTATGCTATGAGATTTCTGCGAGTTTGTATGATACA 1917
Oy 1189 CGGAATAATGGAGATAGACTTGGTCTTGAAGAGGGACTCTGTGACCTGCTTATAGGC 1248
Db 1918 CGGAATAATGGAGATAGACTTGGTCTTGAAGAGGGACTCTGTGACCTGCTTATAGGC 1977
Oy 1249 AAAATTGATGAATCCGTTATGTTGTTGGCTTGAAGAGTGAAGATTTCAATATAGGAAT 1308
Db 1978 AAAATTGATGAATCCGTTATGTTGTTGGCTTGAAGAGTGAAGATTTCAATATAGGAAT 2037
Oy 1309 CCAGCTATTGAGAGTTGTTTGGAGATTTCAATGATGATGATGATGATGATGATGATGATG 1368
Db 2038 CCAGCTATTGAGAGTTGTTTGGAGATTTCAATGATGATGATGATGATGATGATGATGATG 2097
Oy 1369 GGACTATGCAAAATTTGTGAAAAGCTGGTGGCAGGGGTTGCTTTCTTCTAGGTTCAAGAC 1428

```

```

Db      2098  GGACATATGCAATTTGTTGAAACCTGTTGGACAGGGGTTGCTTTCCATGAGTTCAGAGAC 2157
Qy      1429  ACCAAAGATTAATAATTTAACTCGGGGACTACTATGATGATCCATGAGTTTGGATTAC 1488
Db      2158  ACCAAAGATTAATAATTTAACTCGGGGACTACTATGATGATCCATGAGTTTGGATTAC 2217
Qy      1489  TTGGAAGAGTGAAGTGAAGTTCAGGGTCTCTCTTTAGCTGCTGCAACTATGGCAGG 1548
Db      2218  TTGGAAGAGTGAAGTGAAGTTCAGGGTCTCTCTTTAGCTGCTGCAACTATGGCAGG 2277
Qy      1549  ATTGGAAGCCGAGCATGTTGAAAGCTATGCTATGCGAGCACTGCGAAGATTTTCTTCC 1608
Db      2278  ATTGGAAGCCGAGCATGTTGAAAGCTATGCTATGCGAGCACTGCGAAGATTTTCTTCC 2337
Qy      1609  CGCTATACAGATGAGAACTCGGCTGAACCCAGAGATGTCAGAGACAGTGTATTAGTGA 1668
Db      2338  CGCTATACAGATGAGAACTCGGCTGAACCCAGAGATGTCAGAGACAGTGTATTAGTGA 2397
Qy      1669  GATCCGTGTTGTAACATGTAAGCCGATGATGATGAGCCCTGATGCTTTATTTGCAAGACT 1728
Db      2398  GATCCGTGTTGTAACATGTAAGCCGATGATGATGAGCCCTGATGCTTTATTTGCAAGACT 2457
Qy      1729  GTAAGACCCCTCTGAAACCTTTGAAACTATGATTAATGCAATTCGAGCTGGGCTCTCAGAG 1788
Db      2458  GTAAGACCCCTCTGAAACCTTTGAAACTATGATTAATGCAATTCGAGCTGGGCTCTCAGAG 2517
Qy      1789  AGTAGCCGTTGTAAGAACTACTGTTGAAATGTCCTGTTGATGATGTTAAAGAGCAGAT 1848
Db      2518  AGTAGCCGTTGTAAGAACTACTGTTGAAATGTCCTGTTGATGATGTTAAAGAGCAGAT 2577
Qy      1849  GTGAAGATCTAGCTGCTGTTGTCGCAATTGAGATGATTCACCTGTCAGCCAGAGATAT 1908
Db      2578  GTGAAGATCTAGCTGCTGTTGTCGCAATTGAGATGATTCACCTGTCAGCCAGAGATAT 2637
Qy      1909  TTTCTTAAGCAGCTCATCTTTTCAAGCAGAGATATGTTCTCTTCAATGGAATCTGAT 1968
Db      2638  TTTCTTAAGCAGCTCATCTTTTCAAGCAGAGATATGTTCTCTTCAATGGAATCTGAT 2697
Qy      1969  GTCCGCTACCA----- 1978
Db      2698  GTCCGCTACCAATGATGATTAATGATGATGATGATGATGATGATGATGATGATGATGAT 2757
Qy      1979  ----- 1978
Db      2758  TGCTTGTGTTGAGCTGAAGAACATAGTCCCATTAATACATGTCACCAAGTTGTAC 2817
Qy      1979  ----- 1978
Db      2818  AAGATTAAACAAGTTGCTGAGTAATTTCACTAATTAATGCTGCTGGAATTTTGTATCAA 2877
Qy      1979  ----- 1978
Db      2878  CTGTAAGACAGAAATTAATTTCACTCAATTTCTGTTAGAAATAAGTAGATTAG 2937
Qy      1979  ----- 1979
Db      2938  AGATTGCTTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2997
Qy      1980  AGGCTGCTGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2039
Db      2998  AGGCTGCTGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 3057
Qy      2040  GAATATAGTATCAAGTGGCAGAGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2099
Db      3058  GAATATAGTATCAAGTGGCAGAGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3117
Qy      2100  AGAAATGTTACAG----- 2113
Db      3118  AGAAATGTTACAGAGTGAAGGAAATTAATCTACAAATTCATCAATGTTGTGAAGACTGT 3177
Qy      2114  -----AGTTTGTGATGGC 2128

```

```

Db      3178  TGACATGATTAATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3237
Qy      2129  GAATGCTGAAGATTTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2188
Db      3238  GAATGCTGAAGATTTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 3297
Qy      2189  ATTATACACTGTTGAAACTATCTGTTGACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2248
Db      3298  ATTATACACTGTTGAAACTATCTGTTGACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 3357
Qy      2249  CTCTGTTGGAAGCACTCTGAGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2308
Db      3358  CTCTGTTGGAAGCACTCTGAGAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3417
Qy      2309  ACAATGCTACTGATGTCAGAACCTACACAAAGATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 2368
Db      3418  ACAATGCTACTGATGTCAGAACCTACACAAAGATTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 3477
Qy      2369  GGTGAAATCACTGAAGGCTCTGTTCTTGCATCATTA 2406
Db      3478  GGTGAAATCACTGAAGGCTCTGTTCTTGCATCATTA 3515

RESULT 6
AB016888
LOCUS
DEFINITION
Arabidopsis thaliana genomic DNA, chromosome 5, PL clone:MDH9.
ACCESSION
AB016888 BAO00015
VERSION
AB016888.1 GI:3449329
KEYWORDS
SOURCE
ORGANISM
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1
REFERENCE
AUTHORS
Aasamizu, E., Sato, S., Kaneko, T., Nakamura, Y., Kotani, H., Miyajima, N.
and Tabata, S.
TITLE
Structural analysis of Arabidopsis thaliana chromosome 5. VIII.
Sequence features of the regions of 1,081,958 bp covered by
seveneen physically assigned P1 and TAC clones
JOURNAL
DNA Res. 5 (6), 379-391 (1998)
PUBMED
10048488
REFERENCE
AUTHORS
Nakamura, Y.
TITLE
Direct Submission
JOURNAL
Submitted (18-AUG-1998) Yasukazu Nakamura, Kazusa DNA Research
Institute, Department of Plant Gene Research, 1532-3, Yana,
Kisarazu, Chiba 292-0812, Japan (E-mail: ynakamu@kazusa.or.jp,
Tel:81-438-52-3935, Fax:81-438-52-3934)
ADDRESS
For the latest information on annotation of this clone, please see
http://www.kazusa.or.jp/kaos/cgi-bin/sgd.graph.cgi?c=MDH9
Genes with similarity to proteins in the databases are described in
'product' or 'note' qualifiers. Genes that have no significant
protein similarity are described as 'unknown protein'.
The software programs used to predict genes include: Graal
(Informatics Group, Oak Ridge National Laboratory,
http://compbio.ornl.gov/Graal-1.3//),
GENSCAN (Chris Bugge, MIT, http://CCR-081.mit.edu/GENSCAN.html),
NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of
Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and
SplicePredictor (Volker Brendel, Stanford University,
http://gremind.zool.jacstate.edu/cgi-bin/sp.cgi).
Genes encoding tRNAs are predicted by tRNAscan-SE
(Sean Eddy, Washington University School of Medicine, St. Louis,
http://genome.wustl.edu/eddy/tRNAscan-SE/).
This sequence may not be the entire insert of this clone. It may be
shorter because we remove overlaps between neighboring submissions.
The 5' clone is K5014 and the 3' clone is K16E1.
location/Qualifiers
1..85791
/organism="Arabidopsis thaliana"

```

/mol_type="genomic DNA"
/db_xref="taxon:3702"
/chromosome="5"
/clone_1fb="Wileui P1"
/ecotype="Columbia"
complement(59. .239)
/note="CDS is reported in Acc# AB203032
gene_id:K5J14.12
unknown protein"
/number=1
/evidence=not experimental
complement(join(2056. .2079,2278. .2393,2991. .3633))
/note="contains similarity to unknown protein
gene_id:MDH9.1
pir||T04864"
/pseudo
/codon_start=1
/evidence=not experimental
join(5739. .6138,6462. .6595)
/note="unnamed protein product; gene_id:MDH9.2
unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAB10474.1"
/db_xref="GI:9759469"
/translation="MVDFEAEEDRKKSSNLSIAKNSVPLFERGRFYEEYSARNE
RLRRKGEAVGVGGTGPVNLGVNHEPMTRKRGTAKESSIKKTIVSVMSSESIPIR
YSLRKEMAKRKPPLPLNVAVASAMKSVTRKEERSLILCTACNTSSGLAFRLITGS
EPGLVHKRRTLPFSLV"
complement(join(6868. .7911,7989. .8414,8555. .8671,
8754. .9137))
/note="gene_id:MDH9.3"
/codon_start=1
/evidence=not experimental
/product="arm repeat containing protein"
/protein_id="BAB10475.1"
/db_xref="GI:9759470"
/translation="MDTDEATGDAENDEVTAEPIIHDEVDAVEIHEEVEKEDD
DCEAGIVSIVSEFLDQINGRYRTOKECFNLVRLKILPFLDEIRVYKAPPEBIM
LNRLAKFLAKKILETSCNGSKITMALDGETMTRPSIYEKILNRVIVYKAPPEBIM
SGDADEIDSLCKOKAKRRITDIDELAVDMVVFSEKTPRNASIITRLAGLE
LQTDIKETETALQSLIQDKGKNIETKQHIIEILNFKKLQGLAEVDILYQPVTK
AITSKTSILPHEFLCPITLIMLDPVIAIGQYKEKSIQKMPDAGHKTCPRKQEL
DHLSTAPNALKNLIMOWCEKKNKFIPEKVSPPDSOXDEVSILVLEASSOLEBO
RSRVOMRLARENPEMCEKKNKFIPEKVSPPDSOXDEVSILVLEASSOLEBO
KLLISNEGAIPVITILENGRREARNSAALFSLSMIDENKVTIGLSNGIPRVYDIL
OHGTRGKDAITLFLNLSANSKRAIDGIVQPLNLLKDKNLGMIDALSTILL
LASHPEGQAIQGLSFIEITLVEFLRQGTPEKCATSVLLEIGSNSSFLIALQFGV
YEYLVETITSGTNRQAKANALIQISKSEQI"
10888. .12579
/note="unnamed protein product; gb|AA07802.1
gene_id:MDH9.4
similar to unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAB10476.1"
/db_xref="GI:9759471"
/translation="WMISEKPLGVRESIRQDELVTTSRRLYKVSQKLKKIKHTEV
VEDEIARGAVNCLISVGCVADTGBDFEDSSNRKSSASEGKGLMTTGTETRL
DCFSYGRERFWKKNRRYIADSGODYRKHVYLPDILMCLMRPLTSLNAHLVCK
KQSMANTQRFLOMRREGSFQTPWLFPAALKDQSGSDIGYDVSQDKMRITETDL
KGRFMYSVTSLIHEEIIYVIGRSMDRNSFKSRGILVFPSPILKAMKILASMRHARSLPI
VGATEVTSFSTQTKONRQDRPHLSRVGSESDYEDPHLSVARQNRNSADONGTK
SHRLIROKDLRLNRSKRFPVLIAGTGGLDEPLDSEIYDSATNTSEMRQPMGF
GVASCGIICNGIFTAISNDKLSGDIIRGFWITITQSPIPRVEHEFPKIVSCNHL
FMLSVMCDEBDGQIGRRNKAVRKLMEIDLVIYLTWTEVSHPDAPMDNATVSDQNI
LMGVEMFKIFQGVLSFFVQCDILTEASWRVSNRQKNTLSCMNTIALHLH"
complement(15087. .16778)
/note="unnamed protein product; gb|AA07802.1
gene_id:MDH9.5
similar to unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAB10477.1"
/db_xref="GI:9759472"
/translation="WMISEKPLGVRESIRQDELVTTSRRLYKVSQKLKKIKHTEV
EDEARGAVNCLISVGCVADTGBDEDEHLSNRKSSASEDKGLMTTGTETRL
DCFSYGRERFWKKNRRYIADSGODYRKHVYLPDILMCLMRPLTSLNAHLVCK
KQSMANTQRFLOMRREGSFQTPWLFPAALKDQSGSDIGYDVSQDKMRITETDL
KGRFMYSVTSLIHEEIIYVIGRSMDRNSFKSRGILVFPSPILKAMKILASMRHARSLPI
VGATEVTSFSTQTKONRQDRPHLSRVGSESDYEDPHLSVARQNRNSADONGTK
SHRLIROKDLRLNRSKRFPVLIAGTGGLDEPLDSEIYDSATNTSEMRQPMGF
GVASCGIICNGIFTAISNDKLSGDIIRGFWITITQSPIPRVEHEFPKIVSCNHL
FMLSVMCDEBDGQIGRRNKAVRKLMEIDLVIYLTWTEVSHPDAPMDNATVSDQNI
LMGVEMFKIFQGVLSFFVQCDILTEASWRVSNRQKNTLSCMNTIALHLH"
complement(join(18111. .18193,18281. .18535,18773. .19001,
19088. .19195,19307. .19447,19542. .19952,20129. .20245))
/note="unnamed protein product; gene_id:MDH9.6
similar to unknown protein
sp|P42251"
/codon_start=1
/evidence=not experimental
/protein_id="BAB10478.1"
/db_xref="GI:9759473"
/translation="MNQISFWCSLLTIVYIAAEPTITRABSCANQASAPQPIHMDA
INKPDPOLFTLMDNITGDIRPLTVIGKERTFEPWRNSPRFVSSSEEMKLRYAKAK
ANPGYSRIQRKAVYIGWDDHDYGLNDAKGFDEKRVNOKMLDPLNEPLDSPRKA
GVYASYVYGPENRKVYIYLDTRYHRPILSDSGILSDTQWMDLENELSGRSEITII
GSSVQVSNLSATGTPLPFYMESWGRFPKERSYPLKILADTRDGVIFISGVHGEIT
RYDQSVYPLDYVTSGLINOSVEKVRPRLSRLPLWYTPMTWYINDCKRKECT
YGCQNFPAISIDWMANVYITLERTDVNGDTVLGTNVSLSQGSNSLKDTTTCCK
SQRYCTLEIDPGRYRLAVLIYFTAVLAMALGLIGAVLAITACVYKCDV"
complement(21540. .22097)
/note="unnamed protein product; contains similarity to
calmodulin
gene_id:MDH9.7"
/codon_start=1
/evidence=not experimental
/protein_id="BAB10479.1"
/db_xref="GI:9759474"
/translation="MTLAKNOKSLSRLYKVSSEKSSRNLEDESRTSSNSGSSS
LNVELTVPDYMANSQDGTSGRELSVSLIGALISRVVEVTSVDVDDGGFID
FEBRLKMEBGDSDEKRRKIKKAPRMVMEGSEFTTAASRLTSLRSGSCYVDAC
KVMIRGFQDNDGDLSTDFEFLVMNR"
join(24090. .24289,24410. .24614,24701. .24945,25115. .25313,
25632. .25679,25821. .25928,26039. .26227,26314. .26613,
26730. .26831,27166. .27301,27487. .27711,27919. .28122,
28213. .28380,28646. .28795,28889. .29026,29204. .29431,
29519. .29671,29784. .29910,30115. .30241,30554. .30674,
30763. .30843,30932. .31033,31112. .31262,31338. .31429)
/note="gene_id:MDH9.8"
/codon_start=1
/evidence=not experimental
/product="pif1lysin"
/protein_id="BAB10480.1"
/db_xref="GI:9759475"
/translation="MASSSSSIFGVSPFLAPFNSGDSRRSHYLSNRNKVAFNPS
SPRLTPRVRVEAPSLIPYNGLMAAONSKGRKKNIVSGKEKGTLSQGRFCFL
CKRNOAGIRRALPSAFVDRTAFLSLSLSLRSKSHOYINATGRIPEHAPAAWD
GIYAEQRDLDLPEIDLSALEAFGLCELSHPRKLRHQKNGRIYILPRKVPNPF
EAMNEVAGSIDEBEDQGLAHMTEHAFJSSKRGKRLGILGASNAYTPHRTVHI
HSPLHTTDSDDLEPVSVDLNLAEIAPFKPLSSVEKERAIILSELMMNTIEYRVC
QLQHLQSENKLDLPPRPIGAEQIKKMDVRIKRFHRRVAPNATVITYGDDINIR
IVNHIENVFKNGLDNSTSSPSPGAFGAMANFLVPLPAGLGATSENEKNTAOS
KMLIKREHAIIRPVEHNSI.PGTSVLD.KPQIPFHEILQNPALIMPCKIPVSKVQTS
DLYNRIMKPIFLSAHRIYTRKSSNPPTSVEISLDRGEGCTGYITLVAPON
MUNAIVYAVOERRLKRGVYTRGELITRYMDALKDSHLLAAMDVNSVYDLDITMS
DALSHYWDQDQHEITLVAAGVYTLAEVTVGAKVLETSDFGRPAAPLPAALVAV
PTKVAIVGVESDFNISPDIISVYKSGIAPLEABELLEVPKELISQDLKEITLQR
NPGVPI.PGSGILTLHDKETGITQLRSNGILAVYKKSSTESAGVRLVVGGRALR
TSQSGAVVGVTRTISEGAVGDSRQVLFVNNHINSLSSTESERFIEMPEFTLR
DNMGQAAPOLLHVLTERSVALLEDAFDARQOLYISYPSISLISERATRAHLLMTAMNG
DERFVETPTKSLQSNLESYKADVAMSHFVUDNMEVSIYVGPSEBEITRCLIDLYTYK
ASHDSKPPGSEPIELFRQYAGIQOVFLKTDTRCAITAGAPRKGFTYDGDIL
FQVSXKL.PVAHDGLKSEBQLBEGDEILQKLLRAHPLFGVYTMGLLAEIINSRLFTT

CDS

VRDLSGLTYDVSFEINLFDRLNGWYIVSVTSPEKVKYKAVDACKSVLRGLHSNOIAP
RELIDRAKRTLLMRHEAEISKSNAYMLNLALQASVPEKSEIKELIEALYSASIED
IYLAWNRKVEDSDLSYSCIGLAGAQAQBEIIVLSEEBEPDVFSGVAVGKSGSMTER
PTT"
Complement {join(33251..33326,33512..33663,33744..33797,
33888..33968,34143..34265,34359..34768,34873..35052,
35052..35052)}

Query Match 71.4%; Score 1717.6; DB 15; Length 85791;

Best Local Similarity 79.1%; Pred. No. 0;

Matches 2402; Conservative 0; Mismatches 4; Indels 632; Gaps 5;

QY 1 ATGGAAGCTTGAAGTCTGGGCAATGGTCTCTCTCCCATTCATTTATGCCGATTACA 60
DB 64077 ATGGAAGCTTGAAGTCTGGGCAATGGTCTCTCTCCCATTCATTTATGCCGATTACA 64136
QY 61 CCGGCGAGCAAAAGCTCGAGAGTAGCCAAACACCTGTACAACTATCTGCTCGCGAGC 120
DB 64137 CCGGCGAGCAAAAGCTCGAGAGTAGCCAAACACCTGTACAACTATCTGCTCGCGAGC 64196
QY 121 AAATGGGCGAGCCGTCTTCTCCGACTTCAATTTCACTCGATTTCTCTCTCTCTCC 180
DB 64197 AAATGGGCGAGCCGTCTTCTCCGACTTCAATTTCACTCGATTTCTCTCTCTCTCC 64256
QY 181 TTGCGCACCGGCAACCAACCGGCACTCTGCTCTCTGCGACCATTTATTTGATGTCTCC 240
DB 64257 TTGCGCACCGGCAACCAACCGGCACTCTGCTCTCTGCGACCATTTATTTGATGTCTCC 64316
QY 241 GAAGCGCAGTCCCGCATCCCATTTGATTTCTACAGATTTAGAGGCTCAAAACATTTTC 300
DB 64317 GAAGCGCAGTCCCGCATCCCATTTGATTTCTACAGATTTAGAGGCTCAAAACATTTTC 64376
QY 301 TTAAACGATGAATAGAAAGAGCATTTGAAAGTAGGTTTGAAACCGCGCAATTCGAT 360
DB 64377 TTAAACGATGAATAGAAAGAGCATTTGAAAGTAGGTTTGAAACCGCGCAATTCGAT 64436
QY 361 TTACGCGACGAGCTTTTATACGCCGAGACAGATTTCTTCAAGCTCTTGGGAACTCTG 420
DB 64437 TTACGCGACGAGCTTTTATACGCCGAGACAGATTTCTTCAAGCTCTTGGGAACTCTG 64496
QY 421 TCTAATCTCGGTCTAGAAAGAGATGACATGAAAGGCTCTTGTATGATGAAGAAGCTACA 480
DB 64497 TCTAATCTCGGTCTAGAAAGAGATGACATGAAAGGCTCTTGTATGATGAAGAAGCTACA 64556
QY 481 GTCATCACTGATGTTCTTGGGAT----- 504
DB 64557 GTCATCACTGATGTTCTTGGGATAGGTAATTTGCAATTAATAAGTTCTTCT 64616
QY 505 -----AGGTTCTTGGGCT 518
DB 64617 GTTTTAATTTCAATGAATTGATTAAGAAAGAACTTTATCTAATGAAGGTTCTGCTGC 64676
QY 519 TCTCTGTGATTAAGAAAGGAGTGTGAGACTGAGATAGTCTTGGGTTGGAGGCTCT 578
DB 64677 TCTCTGTGATTAAGAAAGGAGTGTGAGACTGAGATAGTCTTGGGTTGGAGGCTCT 64736
QY 579 GCTTAAGAGAGGTTGCTTAAGTCTTTAAGCAAGATGTGTTTATGATTAAGCGCTTGC 638
DB 64737 GCTTAAGAGAGGTTGCTTAAGTCTTTAAGCAAGATGTGTTTATGATTAAGCGCTTGC 64796
QY 639 GTTTCTCATGTTCTGAGAGGATGCTATGAGCATTTGATTTATTTATTAATCTGTTA 698
DB 64797 GTTTCTCATGTTCTGAGAGGATGCTATGAGCATTTGATTTATTTATTAATCTGTTA 64856
QY 699 TGAAGTTTGTGAGAGCTTTGAAGCTTTTAC----- 730
DB 64857 TGAAGTTTGTGAGAGCTTTGAAGCTTTTACAGGATGTTGACTTGGCTTGTAAATTTG 64916
QY 731 -----A 731
DB 64917 ACGAGCTTGCTTATTAAGAACTTTCTGATTTGATACCTTTGATTAAGAGCTTGTGTA 64976
QY 732 GGAGGAAGAGAGATGAGCTTGTGACCGGATTTAGTGCAAAATTTGATGAGACTTTTGA 791
DB 732 GGAGGAAGAGAGATGAGCTTGTGACCGGATTTAGTGCAAAATTTGATGAGACTTTTGA 791

DB 64977 GGAAGAAAGAGCAAGTACCTTGCAACCGATTTACGTGCAAAATTTGATGAGACTTTGA 65036
QY 792 AGAGATCACTCCGCGTATGTTCTTGAGACTTGGCTTACCGCTTGGTATGATTAACG 851
DB 65037 AGAGATCACTCCGCGTATGTTCTTGAGACTTGGCTTACCGCTTGGTATGATTAACG 65096
QY 852 TGCAGAAAGACTAATGTTTAAAGCGGTGCGGAATATTTTGTGCTGTGTGAGAGAGG 911
DB 65097 TGCAGAAAGACTAATGTTTAAAGCGGTGCGGAATATTTTGTGCTGTGTGAGAGAGG 65156
QY 912 TGGAGCATCAAGCTTGTGGGGGTTTGAACCGGAGAAATTTATGAAATGAGGCTTTT 971
DB 65157 TGGAGCATCAAGCTTGTGGGGGTTTGAACCGGAGAAATTTATGAAATGAGGCTTTT 65216
QY 972 ACGAATGACAGCTGCTGAGC----- 991
DB 65217 ACGAATGACAGCTGAGAGAGGATACGTTAGATCTTTTATTTTATTTCTTTAGCA 65276
QY 992 -----AGGTTGATCTTTTGTGA 1008
DB 65277 TGATATAACTTTAGGTTTCTCAATTTATGATATGTTGTGATAGGTTGATCTTTTGTGA 65336
QY 1009 GCTAACCCCAAGCAATATTTCCAGAGAGTCAATTTGAAGTTTACGAAGTTGACTGCTCT 1068
DB 65337 GCTAACCCCAAGCAATATTTCCAGAGAGTCAATTTGAAGTTTACGAAGTTGACTGCTCT 65396
QY 1069 GTGGCTCAAGCTTTATTTAGTAAGAAAGCCACACTTTTACAGAGTGTGATTAAGCAATTC 1128
DB 65397 GTGGCTCAAGCTTTATTTAGTAAGAAAGCCACACTTTTACAGAGTGTGATTAAGCAATTC 65456
QY 1129 CAGCACTTACAGAGGCTAAGTATGCTATGAGAGATTTCTGCGATGTTGATGATACA 1188
DB 65457 CAGCACTTACAGAGGCTAAGTATGCTATGAGAGATTTCTGCGATGTTGATGATACA 65516
QY 1189 CGGAATATTTGGGAGATGAGCTTCCGTGTAAGAAAGGGAATCTGTGACGCTTATAGGC 1248
DB 65517 CGGAATATTTGGGAGATGAGCTTCCGTGTAAGAAAGGGAATCTGTGACGCTTATAGGC 65576
QY 1249 AAGGTTGATATGCTGATATGTTGGGCTTGTAGACAGTGAAGATTCACAAATATAGGAAT 1308
DB 65577 AAGGTTGATATGCTGATATGTTGGGCTTGTAGACAGTGAAGATTCACAAATATAGGAAT 65636
QY 1309 CCAGCTATTGTGAGATTGTTTGTGAGAAATTCAAATGTGATGACAAATGATCTTCTCT 1368
DB 65637 CCAGCTATTGTGAGATTGTTTGTGAGAAATTCAAATGTGATGACAAATGATCTTCTCT 65696
QY 1369 GGAATATGAAATTTGTGAAACCTGTGTGGCAGGGTGTCTTCTAGATTCAAGAGAC 1428
DB 65697 GGAATATGAAATTTGTGAAACCTGTGTGGCAGGGTGTCTTCTAGATTCAAGAGAC 65756
QY 1429 ACCAAAGATATAAATTTTAACTCGGGGACTACTATGATGATCTTGTGATTAAC 1488
DB 65757 ACCAAAGATATAAATTTTAACTCGGGGACTACTATGATGATCTTGTGATTAAC 65816
QY 1489 TTGGAAGAGTGAAGGATGTTCAAGGTTCTCTTTAGCTCTGCTGCACTATGCAAGG 1548
DB 65817 TTGGAAGAGTGAAGGATGTTCAAGGTTCTCTTTAGCTCTGCTGCACTATGCAAGG 65876
QY 1549 ATTGAGACCGAGATGGAAGGCTAGTGTATGAGGACATGTGCAAGAAAGTTTCTCTCC 1608
DB 65877 ATTGAGACCGAGATGGAAGGCTAGTGTATGAGGACATGTGCAAGAAAGTTTCTCTCC 65936
QY 1609 CGCTATACAGATAGAAACTCGGCTGAACCCAGAGATGTGCAAGACAGTGTATGATGA 1668
DB 65937 CGCTATACAGATAGAAACTCGGCTGAACCCAGAGATGTGCAAGACAGTGTATGATGA 65996
QY 1669 GATCTGTGTGTAACAATGTAGAGCGCTGATGTGAGCTGTGTCTTTATTTGCAAGAGCT 1728
DB 65997 GATCTGTGTGTAACAATGTAGAGCGCTGATGTGAGCTGTGTCTTTATTTGCAAGAGCT 66056
QY 1729 GTAAGACCTCTGAAACTTTGAAACTAATGATTAATGCAATTTGAGAGCTGGGGTCTCAGAG 1788
DB 66057 GTAAGACCTCTGAAACTTTGAAACTAATGATTAATGCAATTTGAGAGCTGGGGTCTCAGAG 66116


```
Dh 132882 ATTGAGTTCGTGTCGAATATGCTTCTGATGACTCACTTGATCATGTAGAGAGAGGG 132941
Qy 743 CAAGTAGCCTTGACACCGATTTAGCGTGCACAAATTGATGAGACTTTGAGAGATCATCTC 802
Db 132942 CAGGAGCCTTACACCGAGTTTACAAACACAAATTGATGAGACTTTGAGAGATCACCC 133001
Qy 803 CGCGTTATGCTTGGAGCTACTTGGCTTACCGCTTGATGATTTACCTGCGAAAGAC 862
Db 133002 CACGTTGTTGTTATTAACCTTTAGCCTTGCCCTTGATGATGAACATGAGCGCGAGAG 133061
Qy 863 TTAATGTTTAAAGCGGTGCGGGAATTTTGTGCTGTTGGAGAGGTGAGCATCAG 922
Db 133062 AAGAAAGTCTCCAAAGTCTCCGCAATCTGTGGCAGTTGAGAGTGGAGAGCAGCAG 133121
Qy 923 CTCTGTTGGGGGTTTGAACCGGTGAGAAATTATGATGAGGCGTTTTCAGATGACAG 982
Db 133122 CAATTGCTGGAGTTTCACACGTGAGATTTATGAAATGAGGCAATTCCTGCATATGAAG 133181
Qy 983 CTGCTGAGC----- 991
Db 133182 CTGCCGAAACAGGTTCATAGAAATATATGTCATGTTTATTTTATTTTCTCCGC 133241
Qy 992 ----- 991
Db 133242 CCCCCCTAGTTTGTGTTATGAGAGTTTACATTTATCTTTTCAATTCCTTTTATC 133301
Qy 992 -----AGTTGATCTTTTGTAGCTACCCCGAACAAATTCGACAGAG 1034
Db 133302 TCCAACTCTTTTTTTAAAGTGCACCTTTTGTAGCAACCTTACCAATATTCAGCTGA 133361
Qy 1035 GTCATTTGAAAGTTTACGAAGTTGACCTTGCTGTGGCTCAAGCTTTTATTTGTAGAA 1094
Db 133362 AAGTTTGAAGGCTATGAGGGTTGCACTTGATGTTGTCAGAAAGCCTTTGTAGTAAAA 133421
Qy 1095 GCCACACCTTTTACGAGATGCTGATAGAAATTCAGCAACTTCGACAGCTTAAAGTAT 1154
Db 133422 GCCACATCTTATCCAAAGATGCTGATTAATTTATTCATCAACTTCAACAACTTAAAGTAC 133481
Qy 1155 GGCTATGAGATTCCTGCGATGTTGATGATAC-----ACGGAATTAATGGAGATGAG 1208
Db 133482 AAATATGAGAAATGCTCCCTCTGTTTATATCTCCATGAGATGAGAAAGAAAGTTGA 133541
Qy 1209 CTTGCGTCTAGAAAGGGGACTGTGACCTGCTTATAGCAAAAGTTGAAATGCCGAT 1268
Db 133542 TTTTCATTAGAAAGGGGCTGTGTCACCTGCTGTGGGGAGCTTATCAATGTCGATC 133601
Qy 1269 GTGCTTGGGCTTACAGCATGAGAGATTCACAAATATAGAAATCCAGTATTTGGAGTTGT 13328
Db 133602 ATGTTGGGTCTCGATAGTGAAGCTCACCTTATGAAACCATCTATTTATAGACTTAT 133661
Qy 1329 TTTGAGAAATTCAAATCGTATGACAAATGATATCTCCCTGAGCATATGCAAAATTTGGA 1388
Db 133662 TATGAAAAACGAAAAAGTATGAGAGACATGATCTTCTGTGACCTGTAAATTTGTTGA 133721
Qy 1389 AACTGTTGGCAGAGGGGTGCTTTCTTAGGTTCAAGACACCAAAATTAATAATTT 1446
Db 133722 GACATGTTGATGAGAGGTGTTTCCCTAGGTTTAGAGTACTTAAGAGACAACTTT 133779
```

RESULT 8
AP008208_011/c
WPCOMMENT

Sequence split into 360 fragments LOCUS AP008208 Accession AP008208

Fragment Name	Begin	End
AP008208_000	1	110000
AP008208_001	100001	210000
AP008208_002	200001	310000
AP008208_003	300001	410000
AP008208_004	400001	510000
AP008208_005	500001	610000
AP008208_006	600001	710000
AP008208_007	700001	810000
AP008208_008	800001	910000

AP008208_009	900001	1010000
AP008208_010	1000001	1110000
AP008208_011	1100001	1210000
AP008208_012	1200001	1310000
AP008208_013	1300001	1410000
AP008208_014	1400001	1510000
AP008208_015	1500001	1610000
AP008208_016	1600001	1710000
AP008208_017	1700001	1810000
AP008208_018	1800001	1910000
AP008208_019	1900001	2010000
AP008208_020	2000001	2110000
AP008208_021	2100001	2210000
AP008208_022	2200001	2310000
AP008208_023	2300001	2410000
AP008208_024	2400001	2510000
AP008208_025	2500001	2610000
AP008208_026	2600001	2710000
AP008208_027	2700001	2810000
AP008208_028	2800001	2910000
AP008208_029	2900001	3010000
AP008208_030	3000001	3110000
AP008208_031	3100001	3210000
AP008208_032	3200001	3310000
AP008208_033	3300001	3410000
AP008208_034	3400001	3510000
AP008208_035	3500001	3610000
AP008208_036	3600001	3710000
AP008208_037	3700001	3810000
AP008208_038	3800001	3910000
AP008208_039	3900001	4010000
AP008208_040	4000001	4110000
AP008208_041	4100001	4210000
AP008208_042	4200001	4310000
AP008208_043	4300001	4410000
AP008208_044	4400001	4510000
AP008208_045	4500001	4610000
AP008208_046	4600001	4710000
AP008208_047	4700001	4810000
AP008208_048	4800001	4910000
AP008208_049	4900001	5010000
AP008208_050	5000001	5110000
AP008208_051	5100001	5210000
AP008208_052	5200001	5310000
AP008208_053	5300001	5410000
AP008208_054	5400001	5510000
AP008208_055	5500001	5610000
AP008208_056	5600001	5710000
AP008208_057	5700001	5810000
AP008208_058	5800001	5910000
AP008208_059	5900001	6010000
AP008208_060	6000001	6110000
AP008208_061	6100001	6210000
AP008208_062	6200001	6310000
AP008208_063	6300001	6410000
AP008208_064	6400001	6510000
AP008208_065	6500001	6610000
AP008208_066	6600001	6710000
AP008208_067	6700001	6810000
AP008208_068	6800001	6910000
AP008208_069	6900001	7010000
AP008208_070	7000001	7110000
AP008208_071	7100001	7210000
AP008208_072	7200001	7310000
AP008208_073	7300001	7410000
AP008208_074	7400001	7510000
AP008208_075	7500001	7610000
AP008208_076	7600001	7710000
AP008208_077	7700001	7810000
AP008208_078	7800001	7910000
AP008208_079	7900001	8010000
AP008208_080	8000001	8110000
AP008208_081	8100001	8210000

AP008208_083	8200001	8310000
AP008208_082	8300001	8410000
AP008208_084	8400001	8510000
AP008208_085	8500001	8610000
AP008208_086	8600001	8710000
AP008208_087	8700001	8810000
AP008208_088	8800001	8910000
AP008208_089	8900001	9010000
AP008208_090	9000001	9110000
AP008208_091	9100001	9210000
AP008208_092	9200001	9310000
AP008208_093	9300001	9410000
AP008208_094	9400001	9510000
AP008208_095	9500001	9610000
AP008208_096	9600001	9710000
AP008208_097	9700001	9810000
AP008208_098	9800001	9910000
AP008208_099	9900001	10010000
AP008208_100	10000001	101100000
AP008208_101	10100001	102100000
AP008208_102	10200001	103100000
AP008208_103	10300001	104100000
AP008208_104	10400001	105100000
AP008208_105	10500001	106100000
AP008208_106	10600001	107100000
AP008208_107	10700001	108100000
AP008208_108	10800001	109100000
AP008208_109	10900001	110100000
AP008208_110	11000001	111100000
AP008208_111	11100001	112100000
AP008208_112	11200001	113100000
AP008208_113	11300001	114100000
AP008208_114	11400001	115100000
AP008208_115	11500001	116100000
AP008208_116	11600001	117100000
AP008208_117	11700001	118100000
AP008208_118	11800001	119100000
AP008208_119	11900001	120100000
AP008208_120	12000001	121100000
AP008208_121	12100001	122100000
AP008208_122	12200001	123100000
AP008208_123	12300001	124100000
AP008208_124	12400001	125100000
AP008208_125	12500001	126100000
AP008208_126	12600001	127100000
AP008208_127	12700001	128100000
AP008208_128	12800001	129100000
AP008208_129	12900001	130100000
AP008208_130	13000001	131100000
AP008208_131	13100001	132100000
AP008208_132	13200001	133100000
AP008208_133	13300001	134100000
AP008208_134	13400001	135100000
AP008208_135	13500001	136100000
AP008208_136	13600001	137100000
AP008208_137	13700001	138100000
AP008208_138	13800001	139100000
AP008208_139	13900001	140100000
AP008208_140	14000001	141100000
AP008208_141	14100001	142100000
AP008208_142	14200001	143100000
AP008208_143	14300001	144100000
AP008208_144	14400001	145100000
AP008208_145	14500001	146100000
AP008208_146	14600001	147100000
AP008208_147	14700001	148100000
AP008208_148	14800001	149100000
AP008208_149	14900001	150100000
AP008208_150	15000001	151100000
AP008208_151	15100001	152100000
AP008208_152	15200001	153100000
AP008208_153	15300001	154100000
AP008208_154	15400001	155100000

Query	Best Local Similarity	7.5%; Matches 510; Conservative	Score 180.8; DB 15; Length 110000; Pred. No. 3.1e-37; 0; Mismatches 292; Indels 124; Gaps 3;
AP008208_155	15500001	15610000	
AP008208_156	15600001	15710000	
AP008208_157	15700001	15810000	
AP008208_158	15800001	15910000	
AP008208_159	15900001	16010000	
AP008208_160	16000001	16110000	
AP008208_161	16100001	16210000	
AP008208_162	16200001	16310000	
AP008208_163	16300001	16410000	
AP008208_164	16400001	16510000	
AP008208_165	16500001	16610000	
AP008208_166	16600001	16710000	
AP008208_167	16700001	16810000	
AP008208_168	16800001	16910000	
AP008208_169	16900001	17010000	
AP008208_170	17000001	17110000	
AP008208_171	17100001	17210000	
AP008208_172	17200001	17310000	
AP008208_173	17300001	17410000	
AP008208_174	17400001	17510000	
AP008208_175	17500001	17610000	
AP008208_176	17600001	17710000	
AP008208_177	17700001	17810000	
AP008208_178	17800001	17910000	
AP008208_179	17900001	18010000	
AP008208_180	18000001	18110000	
AP008208_181	18100001	18210000	
AP008208_182	18200001	18310000	
AP008208_183	18300001	18410000	
AP008208_184	18400001	18510000	
AP008208_185	18500001	18610000	
AP008208_186	18600001	18710000	
AP008208_187	18700001	18810000	
AP008208_188	18800001	18910000	
AP008208_189	18900001	19010000	
AP008208_190	19000001	19110000	
AP008208_191	19100001	19210000	
AP008208_192	19200001	19310000	
AP008208_193	19300001	19410000	
AP008208_194	19400001	19510000	
AP008208_195	19500001	19610000	
AP008208_196	19600001	19710000	

```

Db      80900 ATGATATTCCTTTTCACAAAAACCGAATAGCATTCCTCCTCGATAGGTTTGAATTTACAT 80841
Qy      1054 GTTGCACTGCTCTTGCTGGCTCAAGCTTTTATTTGTTAAGAACCCACACTTTTACAGAT 1113
Db      80840 GTAGACATTCGATCGATCGCTCAAGCAATTTATAGTTAAAGGCCACATTCATCATGATG 80781
Qy      1114 GCTGTAAGCAATTCACAGCACTTCAGCAGGCTTAAGGTATGATGAGATTCCTCGC 1173
Db      80780 GCGAGTATCTTTTGAACAATCCAGAAAGTTCA-----C 80745
Qy      1174 ATGTTGATATGATACACGGAATATTTGGAGATAGACTTCGCTTGAAGAGGACTCTGT 1233
Db      80744 ATAGGTTCTCATTTATGCTTATGATATATAGATGACCTTGATGGAAGGGCAATTCGTC 80685
Qy      1234 GCACGCTTATATAGCAAAAGTTGATGATGATGCGGTATGCTGTTGGCTTGAAGAGGAT 1293
Db      80684 TCATTGCTAGTGGAGATGTTTGGCAAGTGCAGAAATGTGGCTTGAATGATATAGACT 80625
Qy      1294 TCACAAATATAGGAATCCAGCTATTTGTGAGTTTGTGGAGAAATTCAAATC--GTGAT 1350
Db      80624 TCACCATACAGAGACCCCAAAATTCAGAGTTTATTTGACCACTTAGCATCAGTGAA 80565
Qy      1351 GACATGATGATCTCCCTGGATATGCAATTTGTTGAAACCTGTTGGCAGGGTTGTC 1410
Db      80564 GAGAAATGATCTTCTCCAGGGCTGTGCAAGCTTTTGGAGACTTGTCTTATCTTTGAGGTT 80505
Qy      1411 TTCTCTAGGTTCAAGACACCAAAAGATTAATACTGGGAGCACTACTATGAT 1470
Db      80504 TTCTCTAGGAGACAGATATCTCGGGCATGCACTTCAAGCTTGGAGATTACTACATGAT 80445
Qy      1471 CCTATGTTTGTAGTTACTTGGAAAGAGTGAAGTATGAGGTTCTCTTAACTGCT 1530
Db      80444 CCAGAAATTTTAAAGCTACTTAAGAAAGATGAGGGTGTGGTCTTCAATTTGCTGCT 80385
Qy      1531 GCTGCACTATGAGCAAGATTGGAGC 1556
Db      80384 GCTGCTGCTATTGCAAAACTTGTGTC 80359

```

RESULT 9

```

AP004885/c  AP004885  150462 bp  DNA  linear  PLN 15-SEP-2004
LOCUS      Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 2,
DEFINITION PAC clone:p0575f10.
ACCESSION  AP004885
VERSION    AP004885.3  GI:41053009
KEYWORDS
SOURCE     Oryza sativa (japonica cultivar-group)
ORGANISM   Oryza sativa (japonica cultivar-group)
            Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzeae; Oryza.
REFERENCE  1
AUTHORS    Sasaki, T., Matsumoto, T. and Yamamoto, K.
TITLE      Oryza sativa nipponbare (GAS) genomic DNA, chromosome 2, PAC
            clone:p0575f10
JOURNAL    Published Only in Database (2002)
AUTHORS    2 (bases 1 to 150462)
TITLE      Sasaki, T., Matsumoto, T. and Yamamoto, K.
JOURNAL    Direct Submission
            Submitted (20-MAR-2002) Takuji Sasaki, National Institute of
            Agricultural Sciences, Rice Genome Research Program, Kannondai
            2-1-2, Tsukuba, Ibaraki 305-8602, Japan
            (E-mail:tsasaki@ias.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
            Tel:81-298-38-7441, Fax:81-298-38-7468)
            On Jan 21, 2004 this sequence version replaced gi:38142429.
            Genes were predicted from the integrated results of the following:
            GENSCAN (http://CCR-081.mit.edu/GENSCAN.html), GENESH
            (http://www.softberry.com/), Genemark.hmm
            (http://opal.biology.gatech.edu/Genemark/), Glimmer
            (http://www.tigr.org/tdb/glimmer/glimr_form.html), RiceHMM
            (http://rgp.dna.affrc.go.jp/RiceHMM/), SplicePredictor
            (http://bioinformatics.iastate.edu/cgi-bin/sp.cgi), sim4

```

FEATURES

SOURCE

(<http://glabin.cse.psu.edu/html/docs/sim4.html>), gap2 (<http://www.tigr.org/software/glimmer/>), BLASTN and BLASTX. The genomic sequence was searched against NCBI Nonredundant Protein database, nr (<ftp://ncbi.nlm.nih.gov/blast/db>) and the cDNA sequence database at RGP or DDBJ. Protein homologues of the coding regions were searched against NCBI Nonredundant Protein database with BLASTP. ESTs represent the identified cDNA sequences using BLASTN with the corresponding DDBJ accession no. and RGP clone ID. Full-length cDNAs represent the identified cDNA sequences using BLASTN with the corresponding DDBJ accession no.

A gene with identity or significant homology to a protein is classified based on the protein name to indicate the homology level such as same name, 'putative-' and '-like protein'. A gene without significant homology to any protein but with full-length cDNA or EST homology (covering almost the entire length of partial sequence) is classified as an 'unknown' protein. A gene predicted by two or more gene prediction programs is classified as a 'hypothetical' protein according to INRSP standard. A gene predicted by a single gene prediction program is also classified as a probable 'hypothetical' protein and is included as a miscellaneous feature of the sequence.

The orientation of the sequence is from 5' to 3' of the PAC clone. This sequence of p0575f10 clone has an overlap with OJ020.C02 (DDBJ: AP004078) clone at 5' end and with P0482f12 (DDBJ: AP005311) clone at 3' end. Detailed information on overlap and assembly quality together with annotation of this entry is available at <http://rgp.dna.affrc.go.jp/GenomeSeq.html>.

Location/Qualifiers

```

1..150462
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="2"
/clone="p0575f10"
/complement(4987..8112)
/gene="p0575f10.1"
/complement(join(4987..5412,5520..5679,5764..5914,6096..6247,6738..6859,7864..8112))
/gene="p0575f10.1"
/feature="supported by full-length cDNA(s): AK106553"
/complement(join(5003..5412,5520..5679,5764..5914,6096..6247,6738..6859,7864..8084))
/gene="p0575f10.1"
/feature="supported by full-length cDNA(s): AK073059"
/complement(join(5315..5412,5520..5679,5764..5914,6096..6247,6738..6859,7864..7876))
/gene="p0575f10.1"
/feature="contains EST(s): AU070160(E61534)
/feature="contains EST(s): AK106553, AK073059"
contains full-length cDNA(s): AK106553, AK073059"
/codon_start=1
/product="putative eukaryotic translation initiation factor 3, subunit 1 alpha, 35kDa"
/protein_id="BAD07941.1"
/db_xref="GI:41053010"
/translation="MEDWSEDEPOPAVPSAKAPLXKWADEDEEDVKSWESEE EKRPPEVEKTAPKPSGKAAGKQQAQSSSVVDALDADLALTKLRQRLVSEAD EKSTTELPKSGSEKSLDFIPKPSDPAVEAEIANKLRPEKSPFHWGLKNTVR LMSASLKGADAKDISSIAIANEKKIKAEKAAAGKKQAGKKKHIEKKDDFLPG KGNFDDDEYDFW"
/complement(10186..13754)
/gene="p0575f10.2"
/complement(join(10186..10479,10759..10892,11127..11991,12089..12349,12515..12712,12843..13073,13455..13754))
/feature="start and end point are not identified"
/complement(join(10186..10479,10759..10892,11127..11991,12089..12349,12515..12712,12843..13073,13455..13754))
/gene="p0575f10.2"
/feature="contains EST(s): AU058423(E51136), AU183658(E51136)
/codon_start=1
/product="plastid division protein precursor"
/protein_id="BAD07942.1"

```

COMMENT

[illegible]

mRNA	complement(join(24479..24817,24913..25266,25398..25633,25720..25832,27054..27185,271995..28148))			
	/gene="P0575Fp10.6-1"			
	/note="supported by full-length cDNA(s) : AK121372"			
	complement(join(24491..24817,24913..25266,25398..25633,25720..25832,27054..27185,271995..28147))			
	/gene="P0575Fp10.6-1"			
	/note="supported by full-length cDNA(s) : AK062255"			
gene	complement(24563..28147)			
	/gene="P0575Fp10.6-2"			
	complement(24563..28147)			
misc_feature	/gene="P0575Fp10.6-2"			
	/note="contains full-length cDNA(s) : AK065148"			
	non-coding transcripc			
	probably inactive due to including stop codon(s) in CDS"			
CDS	complement(join(24787..24817,24913..25266,25398..25633,25720..25832,27054..27159))			
	/gene="P0575Fp10.6-1"			
	/note="contains EST(s) : AU031017(B60571), AU093466(B60571)			
	contains full-length cDNA(s) : AK121372,AK062255"			
	/codon_start=1			
	/product="putative arginine/serine-rich splicing factor			
Query Match	7.5%;	Score 180.8;	DB 15;	Length 150462;
Best Local Similarity	55.1%;	Pred. No. 3.2e-37;		
Matches 510;	Conservative	0;	Mismatches 292;	Indels 124; Gaps 3;
QY	731	AGGAGAAAGGAGCAGTATGCTTGACCCGATTTACGTGCACAAATTGATGAGACTTTGG	790	
Db	12351	AGGAAGATGAGACAGCAATCTCGCACTGATCTGCTTTACAGATGATGAAACTCTCG	12239	
QY	791	AAGAGTCACTCCGCGTTATGTCTTGGAGCTTACTTGCGTTACCGCTTGGTATTAACG	850	
Db	12291	AGGAGATTACACCTCGCTGTGTATGGAGCTTCTCCCTTCATTGACACAGAGCATC	12233	
QY	851	CTGGGAAAACCTTAATGTGTTTAAGCGGTGTGCGGAATATTTGTGTCGTGTGGAGAG	910	
Db	12231	ATAGAAAGCCCAAGAAAGGCGCTTCMAAGTGCAGAAACATTTGTGTGAGCGTTGGCAGAG	1217	
QY	911	GTGGAGCATGAGCTCTTGTTGGGGGTTTGACCCGTGGAAGTTTATGATGAGGCGTTT	970	
Db	12171	GAGGTATTTGCTACCGTTTGGAGAGGATTTTCTGTGAAGCTTCATTAAGAGCTTTT	1211	
QY	971	TACGAATGACAGCTGTCTGACAGGTT-----	996	
Db	12111	TGAGGATGACATCAATTGAACAGTTTGTATCATGTTGTTTATGAAGACGTGCTAT	1205	
QY	997	-----	996	
Db	12051	TTATATTGCAGAGCACATCTTTGAAAACATGTTAACTCCCTTCTGTACACCTCCAG	1199	
QY	997	---GATCTTTTGTAGTACCCCAAGAAATATTCACAGACATTTTGAAGTTTGAAGAA	1053	
Db	11991	ATGGAATTTCTTTCAAAAACACGAATTAAGCATTTCTCTAAATGTTTGAATTTACAAT	1193	
QY	1054	GTTGCACTTGCTCTTGTGCTCAAGCTTTTATTGTGTAAAGCACACCTTTTACAGGAT	1113	
Db	11931	GTAGCACTTGACATGTGCGTCAAGCAATTTAATAGTAAAGGCCACAAATTCATCATGANG	1187	
QY	1114	GCTGATTAAGCAATTCACGCACTTCAGCAGGCTTAAGTAAATGCGTATGAGATTCCTGCG	1173	
Db	11871	GCGGATGATCTTTTGGAAACAATCCAGAAAGTTCAA-----C	1183	
QY	1174	ATGTGTATATATACAGGAATTAATTTGGGAAATATGACTTCGGTCTAGAAAGGGGACTCTGT	1233	
Db	11835	ATAGGTTCTATATATGTCTTAATGATATATGAATGAGACTTTCATATGGAAGGGCAATTTGCG	1177	
QY	1234	GCACTGCTTATPAGGCAAGTTGATGTAATGATGCGGTATGTGTGGGCTTACACATGAGAT	1293	
Db	11775	TCAATGCTATTCGAGAAATGTTACCAAGTGCAGATGTGCGTTGGAAATGATATGATGCT	1171	
QY	1294	TCACAATATAGGAATCCAGCTATTTGTGAGATTGTTGTTGGAGAAATTCAAATC---GTGAT	1350	

```

Db      11715 TCACCATACAGAGACCCCAAAATCTAGATTATTGTGACCACTAGCATCAGTGA 11656
Qy      1351 GACATATGATCTCCCTGGAATATGTAAGTGGTGAAGGCTGTC 1410
Db      11655 GAGATGATCTTCTTCAGGCGTGTGCAAGCTTTGGACTTGTGCTTATCTTTAGGTT 11596
Qy      1411 TTTCTAGGTTGAGACACCAAGATMAAAATTTAACTCGGAGACTATGATGAT 1470
Db      11595 TTTCTAGAGACGAGATCTCGGCGATGACATTGACATTGAGATTAATGATGAT 11536
Qy      1471 CCTATGTTTATGTTACTTGAAGAGAGATGTTAGGTTTCTTACTGCT 1530
Db      11535 CCAGAAATTTAAGCTACTTAAGATGAGAGGCTGTGCTTCTCATTTGGCTGCT 11476
Qy      1531 GCTGCACTATGCAAGATTTGAGC 1556
Db      11475 GCTGCTGCTATTCAGAAACTTGCTG 11450

RESULT 10
LOCUS      DQ022571          550 bp      DNA      linear      STS 28-MAY-2005
DEFINITION Beta vulgaris chromosome 1, sequence tagged site.
ACCESSION  DQ022571
VERSION     DQ022571.1 GI:66394762
KEYWORDS   STS.
SOURCE      Beta vulgaris
ORGANISM   Beta vulgaris
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            Caryophyllales; Amaranthaceae; Beta.
REFERENCE   1 (bases 1 to 550)
AUTHORS     Priesen,T.L., Weiland,J.J., Aasheim,M.L., Hunger,S., Borchardt,D.C.
            and Lewellen,R.T.
TITLE       Identification of a SCAR marker associated with Bm the Beet mosaic
            virus resistance gene on chromosome 1
            Unpublished
            2 (bases 1 to 550)
            Weiland,J.J. and Priesen,T.L.
REFERENCE   Direct Submission
            Submitted (29-APR-2005) Sugarbeet and Potato Research
            USDA-Agricultural Research Service, 1307 18th St. N, Fargo, ND
            58105, USA
FEATURES
    source             Location/Qualifiers
                        1..550
                        /organism="Beta vulgaris"
                        /mol_type="genomic DNA"
                        /culivar="C719"
                        /db_xref="taxon:161934"
                        /chromosome="1"
                        /note="fwd primer seq: GACTGGAGTCGTAAAGCACTGT;
                        rev primer seq: GGAAGCATTTCACTCTTTATGTT;
                        fwd primer_name: Rbm05fwd; rev primer_name: Rbm05rev"
                        1..550
                        /note="SCAR marker associated with resistance to Beet
                        mosaic virus"

ORIGIN
Query Match      6.9%; Score 165.4; DB 10; Length 550;
Beet Local Similarity 74.6%; Pred.No.2.5e-33;
Matches 208; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

Qy      718 TTGAAGCTTTTACAGAGAGAGCACTGACCTTGACCGGATTTAGTCACAAAT 777
Db      102 TTCACTATTTTGTAGAGAGAGGCTTCAATAGCTTTGCCCGGATTTACAGAGAGTT 161
Qy      778 GATGAGACTTTTGAGAGATCACTCGCGTTATGCTTTGAGACTACTTGCTTACCGCTT 837
Db      162 GACGAGACATTTGAGAGAGATCAGACCCCGTTATGCTTGAGACTTTTGCTTGCCTTC 221
Qy      838 GGTGATGATTTAGCGCTCCGAAAGACTTAATGTGTTTAAAGCGGTGCGGAATTTTTCGCG 897
Db      222 AGTGATGAATTTTCGAAAGAAAGATGAAGGCTTCACGCTGTACGCAATATATTTATGG 281

```

```

Qy      898 TCTGTTGAGAGAGGTGAGAGATCAGCTCTTGTGGGGTTTGAACCCGTGAAATTATG 957
Db      282 TCTGTTGAGAGAGGTGTGCGCAGACTATAGCTGTGTTTACTGCTGAGGATTTGATG 341
Qy      958 AATGAGCGCTTTTTCAGAAATGACAGCTGCTGAGCAGGTT 996
Db      342 AATGAGCGCTTTTTCAGAAATGACAGCTGAGCAGGTT 380

RESULT 11
LOCUS      AC155597          184752 bp      DNA      linear      HTG 25-JAN-2005
DEFINITION Zea mays strain B73 clone ZMMB8c0216K08, *** SEQUENCING IN PROGRESS
            **, 26 unordered pieces.
ACCESSION  AC155597
VERSION     AC155597.2 GI:58082456
KEYWORDS   HTG; HTGS_PHASE1.
SOURCE      Zea mays
ORGANISM   Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoidae; Andropogoneae; Zea.
            1 (bases 1 to 184752)
            Chan,A.P., Perlea,G., Zheng,L., Cheung,F., Lee,D., Koo,H.,
            Uteerback,T.R., Feldblyum,T.V., Rabinowicz,P., Fraser,C.M.,
            Schubert,K., SamMiguel,P., Ma,J., Pontaroli,A.C., Rohlfing,T.,
            Budiman,A., Bedell,J.A., Lakey,N., Barbazuk,B., Bennetzen,J. and
            Quackenbush,J.
            Consortium for Maize Genomics - BAC skim sequencing and assembly
            Unpublished
            2 (bases 1 to 184752)
            Chan,A.P., Perlea,G., Zheng,L., Cheung,F., Lee,D., Koo,H.,
            Uteerback,T.R., Feldblyum,T.V., Rabinowicz,P., Fraser,C.M.,
            Schubert,K., SamMiguel,P., Ma,J., Pontaroli,A.C., Rohlfing,T.,
            Budiman,A., Bedell,J.A., Lakey,N., Barbazuk,B., Bennetzen,J. and
            Quackenbush,J.
            Direct Submission
            Submitted (15-JAN-2005) The Institute for Genomic Research (TIGR),
            9712 Medical Center Dr, Rockville, MD 20850
            3 (bases 1 to 184752)
            Chan,A.P., Perlea,G., Zheng,L., Cheung,F., Lee,D., Koo,H.,
            Uteerback,T.R., Feldblyum,T.V., Rabinowicz,P., Fraser,C.M.,
            Schubert,K., SamMiguel,P., Ma,J., Pontaroli,A.C., Rohlfing,T.,
            Budiman,A., Bedell,J.A., Lakey,N., Barbazuk,B., Bennetzen,J. and
            Quackenbush,J.
            Direct Submission
            Submitted (25-JAN-2005) The Institute for Genomic Research (TIGR),
            9712 Medical Center Dr, Rockville, MD 20850
            On Jan 25, 2005 this sequence version replaced gi:57863118.
            ----- Trace submission
            Center name: TIGR
            Seq_id_id: ZGGO
            ----- Project information
            Web site: http://www.tigr.org/tdb/cgi/maize/
            Contact: maize@tigr.org

* NOTE: This is a 'working draft' sequence. It currently
* consists of 26 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1
* 2210: contig of 2210 bp in length
* 2310: gap of unknown length
* 2311
* 2311
* 9941: contig of 7631 bp in length
* 10041: gap of unknown length
* 10042
* 16745: contig of 6704 bp in length
* 16746
* 16845: gap of unknown length
* 32394: contig of 15549 bp in length
* 32395
* 32494: gap of unknown length

```

```
* 32495 35494: contig of 3000 bp in length
* 35495 35594: gap of unknown length
* 35595 40051: contig of 4457 bp in length
* 40052 40151: gap of unknown length
* 40152 41379: contig of 1228 bp in length
* 41380 41479: gap of unknown length
* 41480 44517: contig of 3038 bp in length
* 44518 44564: gap of unknown length
* 44565 45664: gap of unknown length
* 45665 49440: contig of 3776 bp in length
* 49441 49540: gap of unknown length
* 49541 52698: contig of 318 bp in length
* 52699 52798: gap of unknown length
* 52799 55964: contig of 3166 bp in length
* 55965 56064: gap of unknown length
* 56065 59236: contig of 3172 bp in length
* 59237 59336: gap of unknown length
* 59337 70052: contig of 10716 bp in length
* 70053 70152: gap of unknown length
* 70153 84603: contig of 14451 bp in length
* 84604 84703: gap of unknown length
* 84704 86418: contig of 1715 bp in length
* 86419 86518: gap of unknown length
* 86519 88069: contig of 1551 bp in length
* 88070 88169: gap of unknown length
* 88170 89661: contig of 1492 bp in length
* 89662 91357: gap of unknown length
* 91358 91457: contig of 1596 bp in length
* 91458 132727: contig of 41270 bp in length
* 132728 132827: gap of unknown length
* 132828 137468: contig of 4641 bp in length
* 137469 137568: gap of unknown length
* 137569 163843: contig of 26275 bp in length
* 163844 163943: gap of unknown length
* 163944 169858: contig of 5915 bp in length
* 169859 169958: gap of unknown length
* 169959 176100: contig of 6142 bp in length
* 176101 176200: gap of unknown length
* 176201 183079: contig of 6879 bp in length
* 183080 183179: gap of unknown length
* 183180 184752: contig of 1573 bp in length.
* Location/Qualifiers
* 1. 184752
  /organism="Zea mays"
  /mol_type="genomic DNA"
  /strain="B73"
  /db_xref="taxon:4577"
  /clone="ZM58BC0216K08"
  2211..2310
  /estimated_length=unknown
  gap 9942..10041
  /estimated_length=unknown
  gap 16746..16845
  /estimated_length=unknown
  gap 32395..32494
  /estimated_length=unknown
  gap 35495..35594
  /estimated_length=unknown
  gap 40052..40151
  /estimated_length=unknown
  gap 41380..41479
  /estimated_length=unknown
  gap 44518..44617
  /estimated_length=unknown
  gap 45565..45664
  /estimated_length=unknown
  gap 49441..49540
  /estimated_length=unknown
  gap 52699..52798
  /estimated_length=unknown
  gap 55965..56064
```

```
Query Match 6.8%; Score 163.4; DB 14; Length 184752;
Best Local Similarity 54.0%; Pred. No. 1.8e-32;
Matches 506; Conservative 0; Mismatches 296; Indels 135; Gaps 3;

ORIGIN
731 AGAGGAAGGAGCAAGTACGCTGACCGGATTTACGCGCAAAATGATGAGCTTTGG 790
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
108936 AGAGGATGGGCGAAGCAATCTTGACCTGACCTTGCAAGATGATGAACTTTGG 108995
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
791 AAGAGTACCTCCGCGTATGCTTGAGACTTACCGCTTGATGATTTACG 850
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
108996 AGGAGATTACACCTCGTTGTATTTGAGCTTTGCTTCTTCACTGATGAAACATA 109055
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
851 CTGCGAAAAGCTAATGCTTTAAGCGGTGCGGAATATTTGTGCTGTTGAGAG 910
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
109056 AAAATTAACGCCAAGAAAGTCTGCAAGGTGCAAAAACATATTTGTGAGTTCGACGAGG 109115
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
911 GTGAGCATCAGCTCTTGTGCGGGTTGACCGGTGGAAGTTATGATGAGCGTTT 970
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
109116 GTGGATTGCTACTGTTGAGAGAGGATTTCTGTAAGGCTTCATGATGAGGCTTTCT 109175
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
971 TACGAATGACAGCTGTCGAGCAGGTTGATCTTTTGTAGCT----- 1011
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
109176 TGCAGATGACATCAGCTGAGAGGTTGTGTCACAAAGTTAATGACCAACATAGCTGT 109235
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1012 ----- 1011
109236 GTGTGTAGTCTGTCATTTTACTTGAACACAGTTTGAAGTCTTCATCTATTTCTGT 109295
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1012 -----ACCCAGAAGAAATTCGACGACAGCATTTG 1042
109296 TTATATCTCAGATGATTTCTTCTTAACACGGAATAGCATACCTGAATGGTTTG 109355
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1043 AAGTTTGAAGTTGACTTCTCTGTTGCTCAAGCTTTTATTTGTAAGACACACC 1102
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
109356 AGATCTATAGTGTGCACTTCCACATGTGTCAGCAATTTGCAAGTAAAGGCCACAT 109415
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1103 TTTTACAGGATGCTGATTAAGCAATTTCCAGCACTTCAGCAGGCTTAAGGTAATGCTATGG 1162
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
109416 TCATCATGATGGCAGATGAATCTTCCAAACAACTTCAGAAAGTTCAA----- 109461
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1163 AGATTCTCGAATGTTGATGATACAGGAATATTTGGGAGATGAGCTTGCGTGAAGAA 1222
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
109462 -----TATGAGTTCTCATATATCTTATGAAAAAGATGAGACCTCGCATTAGAAA 109511
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```


REFERENCE	1	Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P., Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M., Gebregiorgis,E., Geer,K., Gill,R., Grady,M., Guerra,M., Guevara,M., Gunaratne,P., Haaland,W., Hamill,C., Hamilton,C., Hamilton,K., Harvey,Y., Pavlak,P., Hawes,A., Henderson,N., Hernandez,J., Hernandez,Z., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M., Hollins,B., Howells,S., Huliy,S., Hume,J., Idlebird,D., Jackson,A., Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A., Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C., Kovas,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J., Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J., Lorenshewa,L., Louisedge,H., Lozado,R.J., Lu,X., Ma,J., Maheshwari,M., Mahindaratne,M., Mahmood,M., Malloy,K., Mangum,A., Mangum,B., Mangua,P., Martin,K., Martin,R., Martinez,E., Manning,S., McLeod,M.P., McNeill,T.Z., Meenen,E., Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S., Morgan,M., Morris,K., Morris,S., Mundaasa,M., Murphy,M., Naif,L., Nankervis,C., Neal,D., Newton,N., Ngyuen,N., Norris,S., Nwokwemeh,O., Okwomu,G., Olarnunagoon,A., Pal,S., Parks,K., Paternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C., Popper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.,-L., Puzo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Rellly,B., Rellly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Riz,S.J., Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sison,I., Sitter,C.D., Smaj,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,, Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabory,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Umanai,K., Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wlaczek,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Weinstockern,A., Weiser,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.
TITLE	2	Direct Submission
JOURNAL	2	Unpublished
REFERENCE	3	Submitted (15-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
AUTHORS	3	(bases 1 to 221756)
TITLE	3	Rat Genome Sequencing Consortium.
JOURNAL	3	Direct Submission
REFERENCE	4	Submitted (09-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
AUTHORS	4	On May 9, 2003 this sequence version replaced gi:22771268.
TITLE	4	The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.
COMMENT	4	----- Genome Center Center: Baylor College of Medicine Center code: BCM Web site: http://www.hgsc.bcm.tmc.edu/ Contact: hgsc-help@bcm.tmc.edu ----- Project Information Center project name: GBOM Center clone name: CH230-6B10 ----- Summary Statistics Assembly program: Atlas;

```

Consensus quality: 215154 bases at least Q40
Consensus quality: 216516 bases at least Q30
Consensus quality: 217475 bases at least Q20
Estimated insert size: 232580; sum-of-contigs estimation
Quality coverage: 5x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
  (see http://www.hgsc.bcm.tmc.edu/docs/genbank\_drafc\_data.html).
* NOTE: This is a 'working draft' sequence. It currently
  consists of 1 contigs. Gaps between the contigs
  are represented as runs of N. The order of the pieces
  is believed to be correct as given, however the sizes
  of the gaps between them are based on estimates that have
  provided by the submitter.
* This sequence will be replaced
  by the finished sequence as soon as it is available and
  the accession number will be preserved.
  1 221756: contig of 221756 bp in length.
FEATURES
source
    1..221756
     /organism="Rattus norvegicus"
     /mol_type="genomic DNA"
     /db_xref="taxon:10116"
     /clone="CH230-6B10"
     1..1861
misc_feature
     /note="wgs_end_extension
clone_end:T7"
misc_feature
     /note="clone_boundary
clone_end:T7
site:BCORI
end_sequence:BH357588"
misc_feature
     /note="clone_boundary
clone_end:Sp6
site:BCORI
end_sequence:BH357591"
ORIGIN
Query Match      2.6%; Score 62; DB 14; Length 221756;
Best Local Similarity 43.3%; Pred. No. 7.9e-05;
Matches 290; Conservative 0; Mismatches 380; Indels 0; Gaps 0;
QY      450 TGAAGTCTTCTGATGATGAGAAGACTACACTCATCATGATGTTCTTGGGATTAAGT 509
Db      207991 TGATGTCCTCTGTTGTTGTTGTTGTCCTGCTCATATGCTGTTGTTGTTGTC 20799323
QY      510 TCTCGGGGCTCTCTGTGATTCAGAAAGGTGTGAGACTGAGATAGTTCTTCGGTTGG 569
Db      207931 TGTGCTGCTCTCTGTGCTGTGTTGTTGTCCTGCTACTGCTGTTGTTGTTGTC 2078722
QY      570 TGAGGCTCTGCTTAAGAGAGAGGTGCCAAGCTTAAAGCAAGATNGSTTTTACTTAA 629
Db      207871 TGTGCTGCTCTGTGCTGTTGTTGTTGTTGTCCTGCTGTGCTGTTGTTGTTGTT 2078121
QY      630 GCGCGTTCGTTCTCGATGTCCTCAGAGGATGCTATGAGCATTTGATCCACTGATTTTAA 689
Db      207811 TGTGCTGCTCTGTGTTGTTGCTGTTGTTGCTGCTGTGCTGTTGCTGCTGCTGT 2077525
QY      690 TACTGTTATGAGTTTGTGTAAGAAAGCTTTGAAGCTTTTACAGAGAGAAAGAACATGAG 749
Db      207751 TGTGCTGCTCTGCTGCTGCTGCTGCTGTTGTCCTGCTGCTGCTGTTGTTGCTGT 2076922
QY      750 CCTTGACACCGAATTAAGGTGACCAAAATGATAGACTTTGGAAGAGATCACCCTCGGTTA 809
Db      207691 TGTGCTGCTCTGTGTTGTTGTTGTTGTTGCTGCTGCTGCTGCTGTTGTTGCTGT 2076322
QY      810 TGTCTTGAGACTACTTGAGCTTACCGCTGTGATGATATTACGCTGCGAAAAGACTAATGAG 869
Db      207631 TGTGCTGCTCTGTGTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2075722
QY      870 TTTAAGCGGTGTGCGGAATTTTGTGTGCTGTTGAGAGAGGTGAGCATCAAGCTCTTGT 929

```

Db	207571	TGTTCTGTTGTTGTGTGCTACTGCTGTTGTTCTTGTGTTGTGCTGCTGCTGCTGTTGTG	20751512			
Oy	9350	TGGGGCGTTTGACCCCGTGGAGAGATTATGATATGATAGCGCTTTTACGAATACAGCTCTCTGA	969			
Db	207511	TGTTTGTGTGCTGCCCGCTGAAGTTGCTGTTGTTGTTGTTGTTGCTGCTGCTGCTCTAT	2074522			
Oy	990	GCAGGTGATCTTTTGTGTAGTACCCCAAGCAATATTCAGACAGACTCATTTGAACTTTA	1049			
Db	207451	TGCTGTGTTGTTGTTGTTGTTGCTGCTGCTGCTGCTGCTGTTGCTGTTGTTGTTGTTGCTG	2073922			
Oy	1050	CGAAGTTGCATCTGCTCTTGTGAGCTCACTTTTATTGTTGAAGACCACTTTTACA	1109			
Db	207391	TGCTGCTGCTGTGTGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG	2073322			
Oy	1110	GGATGCTGAT	1119			
Db	207331	TGTTCTGCT	207322			
RESULT 14	AC114193	232974 bp	DNA	linear	HTG 13-MAY-2003	
LOCUS	AC114193/c					
DEFINITION	Rattus norvegicus clone CH230-228E13, *** SEQUENCING IN PROGRESS					
ACCESSION	AC114193					
VERSION	AC114193.5	GI:30579511				
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.					
SOURCE	Rattus norvegicus (Norway rat)					
ORGANISM	Rattus norvegicus					
REFERENCE	Murphy, D., Marie, Metzker, M., Lee, A., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooms, S., Amir, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biwaldo, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Bunay, C., Burch, P., Butrell, K., Calderon, E., Cadena, V., Carter, K., Cavazos, I., Cesar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., David, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Diaper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisi, A., Garcia, R., Garcia, A., Garner, T., Garza, M., Gegregoric, E., Geer, K., Gail, R., Gardy, M., Gaurer, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Haylak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hognes, M., Hollins, B., Howell, S., Huylk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpaty, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorenshewa, L., Louisseged, H., Lozano, R. J., Lu, X., Ma, U., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Manthey, S., McLeod, M. P., McNeill, T. Z., Meenen, B., Milosavljevic, A., Miner, G., Mirza, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidas, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwokenkeme, O., Okwunonu, G., Olarinmugoon, A., Pal, S., Parks, K., Patel, K. S., Paul, H., Perez, A., Perez, L., Pfankuch, C., Plopper, F., Poindexter, A., Popovic, D., Prims, E., Pu, L., L., Puzo, M., Quito, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Kelly, B., Kelly, M., Ren, Y., Reuter, M., Richards, S., Rigs, F., Rives, C., Rodkey, T., Rojias, A., Rose, M., Rose, R., Ruiz, S. U., Sanders, W., Savery, G., Scheer, S., Scott, G., Shatman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajs, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Soosa, J.,					

```

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
1. .232974
/mol type="genomic DNA"

```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 10, 2005, 08:09:29 ; Search time 1361.27 Seconds
(without alignments)
11779.592 Million cell updates/sec

Title: US-10-600-070B-1

Perfect score: 2406

Sequence: 1 atggaagctctgagtcacgt.....gtctgtctctgcataca 2406

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_21: *
1: geneseqn1980s: *
2: geneseqn1990s: *
3: geneseqn2000s: *
4: geneseqn2001as: *
5: geneseqn2001bs: *
6: geneseqn2002as: *
7: geneseqn2002bs: *
8: geneseqn2003as: *
9: geneseqn2003bs: *
10: geneseqn2003cs: *
11: geneseqn2003ds: *
12: geneseqn2004as: *
13: geneseqn2004bs: *
14: geneseqn2005s: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2406	100.0	2406	12	ADJ38129 Arabidops
2	2402.8	99.9	2406	12	ADJ38135 Arabidops
3	2399.6	99.7	2406	12	ADJ38208 Plastid d
4	2399.6	99.7	2637	12	ADJ38210 Plastid d
5	2399.6	99.7	2679	13	ADJ14901 plant CDN
6	1724	71.7	3667	12	ADJ38130 Arabidops
7	1720.8	71.5	3667	12	ADJ38136 Arabidops
8	482	20.0	2283	12	ADJ38212 Plastid d
9	481	20.0	2283	12	ADJ38206 Plastid d
10	328.8	13.7	631	12	ADJ38264 Plastid d
11	278.8	11.6	660	12	ADJ38215 Plastid d
12	224	9.3	537	12	ADJ38223 Plastid d
13	222.4	9.2	552	13	ACN48855 Cotton pr
14	207	8.6	545	12	ADJ38255 Plastid d
15	197.2	8.2	491	12	ADJ38254 Plastid d
16	195.8	8.1	647	12	ADJ38265 Plastid d
17	193.8	8.1	460	12	ADJ38267 Plastid d
18	190	7.9	446	12	ADJ38231 Plastid d
19	186.6	7.8	652	12	ADJ38266 Plastid d

20	178.6	7.4	607	13	ACN48781	ACN48781 Cotton pr
21	175	7.3	439	12	ADP93143	Cotton ex
22	170.4	7.1	608	12	ADJ38217	Plastid d
23	170.2	7.1	563	12	ADJ38258	Plastid d
24	169.8	7.1	622	12	ADJ38226	Plastid d
25	165.4	6.9	527	12	ADJ38214	Plastid d
26	158.6	6.6	871	12	ADJ38232	Plastid d
27	155.4	6.5	307	12	ADJ38218	Plastid d
28	150.4	6.3	418	12	ADJ38224	Plastid d
29	144.6	6.0	653	12	ADJ38228	Plastid d
30	144.2	6.0	479	12	ADJ38230	Plastid d
31	130.8	5.4	420	12	ADJ38256	Plastid d
32	128.6	5.3	480	12	ADJ38225	Plastid d
33	127.6	5.3	535	12	ADJ38229	Plastid d
34	124.2	5.2	360	12	ADJ38259	Plastid d
35	118.8	4.9	336	12	ADJ38222	Plastid d
36	117.8	4.9	309	12	ADJ38221	Plastid d
37	111.4	4.6	540	12	ADJ38233	Plastid d
38	111.2	4.6	604	12	ADJ38227	Plastid d
39	110.6	4.6	606	12	ADJ38257	Plastid d
40	96.8	4.0	295	12	ADJ38213	Plastid d
41	92.6	3.8	187	12	ADJ38216	Plastid d
42	83.2	3.5	300	12	ADJ38260	Plastid d
43	77.6	3.2	400	12	ADJ38263	Plastid d
44	76.8	3.2	416	12	ADJ38219	Plastid d
45	75.2	3.1	549	12	ADJ38261	Plastid d

ALIGNMENTS

RESULT 1
ADJ38129
ID ADJ38129 standard; cDNA; 2406 BP.
XX
AC ADJ38129;
DT 06-MAY-2004 (first entry)
XX
DE Arabidopsis thaliana Arc6-1 cDNA Seg1d1.
XX
KW prokaryotic type; plastid division; Ftn2; ARC6; ARC5; Fzo; plant cell;
KW agronomic; horticultural; crop plant; ornamental plant; woody plant;
KW herbicide target; gene; ss.
XX
OS Arabidopsis thaliana.
XX
FN WO2004001003-A2.
XX
PD 31-DEC-2003.
XX
PP 20-JUN-2003; 2003WO-US019536.
XX
PR 20-JUN-2002; 2002US-0390140P.
PR 09-AUG-2002; 2002US-0402242P.
PR 20-JUN-2003; 2003US-00600070.
XX
PA (UNMS) UNIV MICHIGAN STATE.
XX
PI Osteryoung KW, Vittha S, Kokscharova OA, Gao H;
XX WPI; 2004-082486/08.
XX DR P-PSDB; ADJ38202.
XX
PT New isolated Ftn2, ARC5 and/or Fzo-like nucleic acid sequences, useful
PT for further characterizing plastid division in plant cells, and in
PT varying agronomic and horticultural characteristics of economically
XX important plants.
XX
XX Claim 1; SEQ ID NO 1; 287pp; English.
XX
CC This invention relates to novel prokaryotic type or plastid division and
CC related genes and proteins. In particular, the invention relates to novel

QY 1991 GGGTCAGTCAGAGCTGACGATTCCAGAACCTTCCAGAAATGATGTAGAGCTGCAGAG 2040
DB 1991 GGGTCAGTCAGAGCTGACGATTCCAGAACCTTCCAGAAATGATGTAGAGCTGCAGAG 2040
QY 2041 AATATAGATCCAACTGCGAGAAAGTTAAGTCTGCGCTTTGGCCCTGATCCAGCCGATA 2100
DB 2041 AATATAGATCCAACTGCGAGAAAGTTAAGTCTGCGCTTTGGCCCTGATCCAGCCGATA 2100
QY 2101 GAAATGTTACGAGAGTCTTGGATGGGCGAAATGCTGAAGATTGGACTGACAGACAGCT 2160
DB 2101 GAAATGTTACGAGAGTCTTGGATGGGCGAAATGCTGAAGATTGGACTGACAGACAGCT 2160
QY 2161 GAAATGTCGCGAGCTGGTGGTTATATATATACACTGTTGAAACATATCTGTTGACAGT 2220
DB 2161 GAAATGTCGCGAGCTGGTGGTTATATATATACACTGTTGAAACATATCTGTTGACAGT 2220
QY 2221 GTGACAGTCTCAGCAGATGGAACCGGTGCTGTGGTGAAGCAACTGTGAGAGAGTCTGCT 2280
DB 2221 GTGACAGTCTCAGCAGATGGAACCGGTGCTGTGGTGAAGCAACTGTGAGAGAGTCTGCT 2280
QY 2281 TGTCTATCTGATTTGGTTCATCCAGAAAACATGCTAGTGTGAGAACTACACACATA 2340
DB 2281 TGTCTATCTGATTTGGTTCATCCAGAAAACATGCTAGTGTGAGAACTACACACATA 2340
QY 2341 AGATACGAAAGTTTCTGTGTCAGATCAGGCTGGAAGAAATCACTGAAGGCTGTGTTCTGCA 2400
DB 2341 AGATACGAAAGTTTCTGTGTCAGATCAGGCTGGAAGAAATCACTGAAGGCTGTGTTCTGCA 2400
QY 2401 TCATPAA 2406
DB 2401 TCATPAA 2406

RESULT 2
ID ADJ38135 standard; cDNA; 2406 BP.
XX ADJ38135;
AC
XX
DT 06-MAY-2004 (first entry)

XX Arabidopsis thaliana AetFn2 cDNA SeqID9.
DE
XX
XX prokaryotic type; plastid division; Ftn2; ARC6; ARC5; Fzo; plant cell;
KW agronomic; horticultural; crop plant; ornamental plant; woody plant;
KW herbicide target; gene; ss.
XX
OS Arabidopsis thaliana.
XX
XX WO2004001003-A2.
PN
XX
PD 31-DEC-2003.
PE 20-JUN-2003; 2003WO-US019536.
PF
XX 20-JUN-2002; 2002US-0390140P.
PR 09-AUG-2002; 2002US-0402242P.
PR 20-JUN-2003; 2003US-00600070.
XX
XX (UNMS) UNIV MICHIGAN STATE.
PA
PI Oseeryoung KW, Vittha S, Koksharova OA, Gao H;
XX
XX WPI; 2004-082486/08.
DR
XX
XX P-PSDB; ADJ38203.
XX
XX New isolated Ftn2, ARC5 and/or Fzo-like nucleic acid sequences, useful
PT for further characterizing plastid division in plant cells, and in
PT varying agronomic and horticultural characteristics of economically
XX important plants.
XX
XX Claim 1; SEQ ID NO 9; 287bp; English.
XX

CC This invention relates to novel prokaryotic type or plastid division and
CC related genes and proteins. In particular, the invention relates to novel
CC Ftn2 (ARC6), ARC5 and Fzo-like genes and polypeptides. The methods and
CC compositions of the present invention are useful for further
CC characterizing plastid division in plant cells, in order to vary
CC agronomic and horticultural characteristics of economically important
CC plants, such as crop, ornamental and woody plants. They can also be used
CC as herbicide targets. The present sequence is a cDNA sequence which is
CC related to the invention.

XX
SQ Sequence 2406 BP; 611 A; 491 C; 619 G; 685 T; 0 U; 0 Other;

Query Match 99.9%; Score 2402.8; DB 12; Length 2406;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2404; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGGAAGCTCGAATCAAGTGTGAGTGTCTCCCATTCGAATTATGCGGATTTACA 60
DB 1 ATGGAAGCTCGAATCAAGTGTGAGTGTCTCCCATTCGAATTATGCGGATTTACA 60
QY 61 CCGGCGAGCAAGAAAGCTCCGACGTAGCCACAAACCTCTACAACTATCTGCCGACG 120
DB 61 CCGGCGAGCAAGAAAGCTCCGACGTAGCCACAAACCTCTACAACTATCTGCCGACG 120
QY 121 AATATGCGCGACCGCTCTTCTCTCCGACTCAATTTCACCTCCGATTCCTCTCTCC 180
DB 121 AATATGCGCGACCGCTCTTCTCTCCGACTCAATTTCACCTCCGATTCCTCTCTCC 180
QY 181 TTGCGCACCGGACACACACCGGACCTGCTCTCTGCGACCATATGATGCTGCC 240
DB 181 TTGCGCACCGGACACACACCGGACCTGCTCTCTGCGACCATATGATGCTGCC 240
QY 181 TTGCGCACCGGACACACACCGGACCTGCTCTCTGCGACCATATGATGCTGCC 240
DB 181 TTGCGCACCGGACACACACCGGACCTGCTCTCTGCGACCATATGATGCTGCC 240
QY 241 GAAACCCACGTCGCCATCCCATTTGATTTTACACAGGATTTAGAGGCTCAACACATTT 300
DB 241 GAAACCCACGTCGCCATCCCATTTGATTTTACACAGGATTTAGAGGCTCAACACATTT 300
QY 301 TTAAACGATGGAATCAGAAAGACATTCGAAGCTTGGTGAACCCGCGCAATTCGGT 360
DB 301 TTAAACGATGGAATCAGAAAGACATTCGAAGCTTGGTGAACCCGCGCAATTCGGT 360
QY 361 TTGACCGACGACGCTTTAATATGCGGAGACAGATTTCTCAAGTGTCTTGGAAATCTG 420
DB 361 TTGACCGACGACGCTTTAATATGCGGAGACAGATTTCTCAAGTGTCTTGGAAATCTG 420
QY 421 TCTAATCCTGCGGTCTAGAAAGAGATCAATGAAGTCTTCTTGAATGAAGAAGCTACA 480
DB 421 TCTAATCCTGCGGTCTAGAAAGAGATCAATGAAGTCTTCTTGAATGAAGAAGCTACA 480
QY 481 GTCAATCACTGATGTTCTTGGGATTAAGGTTCTGGGCTCTCTGTGATTCGAAGAGT 540
DB 481 GTCAATCACTGATGTTCTTGGGATTAAGGTTCTGGGCTCTCTGTGATTCGAAGAGT 540
QY 541 GGTGAGACTGAGATGATTTCTTGGGTTGTGAGAGCTCTGCTTAAGAGAGGTTGCTAAG 600
DB 541 GGTGAGACTGAGATGATTTCTTGGGTTGTGAGAGCTCTGCTTAAGAGAGGTTGCTAAG 600
QY 601 TCGTTTAAAGCAAGATGATTTAGTTATGAGCGCTTCTGATGATGAGAGGAT 660
DB 601 TCGTTTAAAGCAAGATGATTTAGTTATGAGCGCTTCTGATGATGAGAGGAT 660
QY 661 GCTATGCGATTGATTCACCTGATTTTATTAATGCTGTTATGAGTTGTTAGAGAAAGCTTGG 720
DB 661 GCTATGCGATTGATTCACCTGATTTTATTAATGCTGTTATGAGTTGTTAGAGAAAGCTTGG 720
QY 721 AAGCTTTTACAGAGGAGAGAGCAAGTACCTGACACCGGATTTACGTCGCAATATGAT 780
DB 721 AAGCTTTTACAGAGGAGAGAGCAAGTACCTGACACCGGATTTACGTCGCAATATGAT 780
QY 781 GAGACTTTGGAAGAGATCACTCCGCTTATGCTTGGAGCTACTTGGCTTATCCGCTTGGT 840
DB 781 GAGACTTTGGAAGAGATCACTCCGCTTATGCTTGGAGCTACTTGGCTTATCCGCTTGGT 840
QY 841 GATGATTAACGCTCGGAAAGACATAATGCTTTAAGCGGTGTGCGGAATATTTTGTGCTT 900

Db 841 GATGATTCGCTGCCAAAAGACTAAATGCTTAAACGGGTGCGGAATATTTTGCTCT 900
Qy 901 GTTGAGAGAGGAGGAGCATCAGCTCTGTGTGGGGGTGTGAACCCGGAAGAATTATGAT 960
Db 901 GTTGAGAGAGGAGGAGCATCAGCTCTGTGTGGGGGTGTGAACCCGGAAGAATTATGAT 960
Qy 961 GAGGCGTTTTTACGAATGACAGCTGCTGACAGTTGATCTTTTGTAGTACCCCAAGC 1020
Db 961 GAGGCGTTTTTATGAAATGACAGCTGCTGACAGTTGATCTTTTGTAGTACCCCAAGC 1020
Qy 1021 AATATTCAGACAGAGTCAATTGGAATTACGAAGTTGCACTTGCTCTGTGGCTCAAGCT 1080
Db 1021 AATATTCAGACAGAGTCAATTGGAATTACGAAGTTGCACTTGCTCTGTGGCTCAAGCT 1080
Qy 1081 TTTATTTGTAAGAAGCCACACTTTTACAGATGCTGATTAAGCAATTCACAGCACTTGCAG 1140
Db 1081 TTTATTTGTAAGAAGCCACACTTTTACAGATGCTGATTAAGCAATTCACAGCACTTGCAG 1140
Qy 1141 CAGGCTAAGGTAATGCTATGAGAAATTCCTGCGATGTTGATGATACACGGAATTAATGG 1200
Db 1141 CAGGCTAAGGTAATGCTATGAGAAATTCCTGCGATGTTGATGATACACGGAATTAATGG 1200
Qy 1201 GAGATAGCTTCGGCTGTGAAAGGGGACTCTGTGCACTGCTTATAGCAAAAGTTGATGA 1260
Db 1201 GAGATAGCTTCGGCTGTGAAAGGGGACTCTGTGCACTGCTTATAGCAAAAGTTGATGA 1260
Qy 1261 TGCCGATGCTGGGCTTGAACAGTGAAGGATTCACAATATAGGAATTCACGCTATGTG 1320
Db 1261 TGCCGATGCTGGGCTTGAACAGTGAAGGATTCACAATATAGGAATTCACGCTATGTG 1320
Qy 1321 GAGTTTGTGTTGAGAAATTCAAATCGTATGATGATGATGATCTCCCTGAAGCTATGCAAA 1380
Db 1321 GAGTTTGTGTTGAGAAATTCAAATCGTATGATGATGATGATGATCTCCCTGAAGCTATGCAAA 1380
Qy 1381 TTTGTTGAAAACCTGCTGAGGAGGCTCTTTCTAGTTCAGAGACACCAAGATATA 1440
Db 1381 TTTGTTGAAAACCTGCTGAGGAGGCTCTTTCTAGTTCAGAGACACCAAGATATA 1440
Qy 1441 AAATTTTAACTCGGGGAACTACTATGATGATCTTATGATGATCTTATGAAAGAGTG 1500
Db 1441 AAATTTTAACTCGGGGAACTACTATGATGATCTTATGATGATCTTATGAAAGAGTG 1500
Qy 1501 GAGGTAGTTCAGGCTCTCTTTAGCTGCTGCTGCACTATGAGAGATTCAGCCGAG 1560
Db 1501 GAGGTAGTTCAGGCTCTCTTTAGCTGCTGCTGCACTATGAGAGATTCAGCCGAG 1560
Qy 1561 CATGTGAAAGCTAGTGTATGACAGGACTGCAAGAAAGTTTCTCCCTATACAGAT 1620
Db 1561 CATGTGAAAGCTAGTGTATGACAGGACTGCAAGAAAGTTTCTCCCTATACAGAT 1620
Qy 1621 AGAAACTCGGCTGAACCCAGAGATGTGCAAGAGACAGTGTATGATGATCTGTTGGT 1680
Db 1621 AGAAACTCGGCTGAACCCAGAGATGTGCAAGAGACAGTGTATGATGATCTGTTGGT 1680
Qy 1681 AACATGTAGGCGCTGATGTGAGCTGTGCTTTATTTAGCAAGAGCTGTAAGACCTCT 1740
Db 1681 AACATGTAGGCGCTGATGTGAGCTGTGCTTTATTTAGCAAGAGCTGTAAGACCTCT 1740
Qy 1741 GAAAACCTTTGAAACTAATGATTAATGCAATTTGAGCTGGGGCTCAGAGAGTGGTTGAT 1800
Db 1741 GAAAACCTTTGAAACTAATGATTAATGCAATTTGAGCTGGGGCTCAGAGAGTGGTTGAT 1800
Qy 1801 GAAACTACTGTTGAAATGTCGGTGTGATGATGTTAAAGAGGCAAGGTGAAGTCTTA 1860
Db 1801 GAAACTACTGTTGAAATGTCGGTGTGATGATGTTAAAGAGGCAAGGTGAAGTCTTA 1860
Qy 1861 GCTGCTGTGTGGCAATTTGAATGATTTCACTGTTCACCGAAGATATTTCTTAAAGC 1920
Db 1861 GCTGCTGTGTGGCAATTTGAATGATTTCACTGTTCACCGAAGATATTTCTTAAAGC 1920
Qy 1921 AGCTCATCTTTTCAACGCAAGGATATGTTCTTCTATGAAATCTGATGTGCTACATA 1980
Db 1921 AGCTCATCTTTTCAACGCAAGGATATGTTCTTCTATGAAATCTGATGTGCTACATA 1980

Db 1921 AGCTCATCTTTTCAACGCAAGGATATGTTCTTCTATGAAATCTGATGTGCTACATA 1980
Qy 1981 GGGTCAGTCAGAGCTGACGATTCAGAAACATCTCCAGAAATGATGCTAGGACTGCAGAG 2040
Db 1981 GGGTCAGTCAGAGCTGACGATTCAGAAACATCTCCAGAAATGATGCTAGGACTGCAGAG 2040
Qy 2041 AATATGATTCAGAGTGGCAGAAAGATTAACTCTGCTTTTGGGCTGATCACCGCATTA 2100
Db 2041 AATATGATTCAGAGTGGCAGAAAGATTAACTCTGCTTTTGGGCTGATCACCGCATTA 2100
Qy 2101 GAAATGTTTACCAAGAGTTTGGATGGGCGAAATGCTGAAAGATTGGACTGACAGACAGCT 2160
Db 2101 GAAATGTTTACCAAGAGTTTGGATGGGCGAAATGCTGAAAGATTGGACTGACAGACAGCT 2160
Qy 2161 GAACTGCGCAGCTTGGGGTGGTTATGATTAATACATGTTGAAACTATCTGTGACAGT 2220
Db 2161 GAACTGCGCAGCTTGGGGTGGTTATGATTAATACATGTTGAAACTATCTGTGACAGT 2220
Qy 2221 GTGACAGTCTCAGAGATGGAACCGTGTCTGTGGAAGCACTCTGAGAGAGTCTGCT 2280
Db 2221 GTGACAGTCTCAGAGATGGAACCGTGTCTGTGGAAGCACTCTGAGAGAGTCTGCT 2280
Qy 2281 TGTCTATCTGATTTGGTTTCATCCGAAACAAATCTACTGATGTCAAGAACTTACACACA 2340
Db 2281 TGTCTATCTGATTTGGTTTCATCCGAAACAAATCTACTGATGTCAAGAACTTACACACA 2340
Qy 2341 AGATACGAAGTTTCTGCTCAAGTCAAGGAGTGAAGAAATCACTGAAGGCTCTGTTCTTGA 2400
Db 2341 AGATACGAAGTTTCTGCTCAAGTCAAGGAGTGAAGAAATCACTGAAGGCTCTGTTCTTGA 2400
Qy 2401 TCATPAA 2406
Db 2401 TCATPAA 2406

RESULT 3
ADJ38208
ID ADJ38208 standard; DNA; 2406 BP.
XX
XX AC ADJ38208;
XX
XX 06-MAY-2004 (first entry)
XX
XX
XX Plastid division-related *Arc6* orthologue gene 2.
XX
XX prokaryotic type; plastid division; *Ftn2*; *ARC6*; *ARC5*; *Fzo*; plant cell;
XX agronomic; horticultural; crop plant; ornamental plant; woody plant;
XX herbicide target; gene; ds.
XX
XX *Arabidopsis thaliana*.
XX
XX
XX WO2004001003-A2.
XX
XX
XX 31-DEC-2003.
XX
XX 20-JUN-2003; 2003WO-US019536.
XX
XX 20-JUN-2002; 2002US-0390140P.
XX 09-AUG-2002; 2002US-0402242P.
XX 20-JUN-2003; 2003US-00600070.
XX
XX (UNMS) UNIV MICHIGAN STATE.
XX
XX
XX OeserYoung KW, Vicha S, Koksharova OA, Gao H;
XX WPI; 2004-082486/08.
XX
XX P-PSDB; ADJ38207.
XX
XX New isolated *Ftn2*, *ARC5* and/or *Fzo*-like nucleic acid sequences, useful
XX for further characterizing plastid division in plant cells, and in
XX varying agronomic and horticultural characteristics of economically
XX important plants.
XX

RS Disclosure; Fig 8; 287bp; English.
 XX This invention relates to novel prokaryotic type or plastid division and
 CC related genes and proteins. In particular, the invention relates to novel
 CC Ftn2 (AFC6), ARCS and Fzo-like genes and polypeptides. The methods and
 CC compositions of the present invention are useful for further
 CC characterizing plastid division in plant cells, in order to vary
 CC agronomic and horticultural characteristics of economically important
 CC plants, such as crop, ornamental and woody plants. They can also be used
 CC as herbicide targets. The present sequence is that of a gene which is
 CC related to the invention.

XX Sequence 2406 BP; 612 A; 493 C; 618 G; 683 T; 0 U; 0 Other;

Query Match 99.7%; Score 2399.6; DB 12; Length 2406;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 2402; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGGAAGCTCTGAGTCAAGTGGGCTGCTCTCCCATTCATTTATGCGATTACCA 60
 DB 1 ATGGAAGCTCTGAGTCAAGTGGGCTGCTCTCCCATTCATTTATGCGATTACCA 60
 QY 61 CCGGGGACGACAAAGCTCCGAGTACGCCAACACCTTACAACTATCTGCTCCGAC 120
 DB 61 CCGGGGACGACAAAGCTCCGAGTACGCCAACACCTTACAACTATCTGCTCCGAC 120
 QY 121 AATATGGCCGACCGCTCTCTCCGACTTCAATTTACCTCGATTCTCTCTCTCC 180
 DB 121 AATATGGCCGACCGCTCTCTCCGACTTCAATTTACCTCGATTCTCTCTCTCC 180
 QY 181 TTGCGCACCGCCACACACCGGCACTCTGCTCTCCGACCATCTATGATGCTGCC 240
 DB 181 TTGCGCACCGCCACACACCGGCACTCTGCTCTCCGACCATCTATGATGCTGCC 240
 QY 241 GAACGCGACGTCGCCATCCCATTTGATTTCTACACAGTATTAAGSACTAAACATTTTC 300
 DB 241 GAACGCGACGTCGCCATCCCATTTGATTTCTACACAGTATTAAGSACTAAACATTTTC 300
 QY 301 TTAAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
 DB 301 TTAAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
 QY 361 TTGACGACGACGCTTTTATTCAGCCGAGACAGATTTCTTAAAGCTGCTTGCAGAACTCTG 420
 DB 361 TTGACGACGACGCTTTTATTCAGCCGAGACAGATTTCTTAAAGCTGCTTGCAGAACTCTG 420
 QY 421 TCTAATCTCGGCTTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
 DB 421 TCTAATCTCGGCTTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
 QY 481 GTGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
 DB 481 GTGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
 QY 541 GGTGAGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
 DB 541 GGTGAGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
 QY 601 TCGTTAAGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
 DB 601 TCGTTAAGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
 QY 661 GGTATGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
 DB 661 GGTATGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
 QY 721 AAGCTTTTACAG 780
 DB 721 AAGCTTTTACAG 780
 QY 781 GAGACTTTGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840
 DB 781 GAGACTTTGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 840

QY 841 GATGATTACGTCGCAAAAAGACTAAATGCTTAAAGCGGTGCGGAATATTTTGTGCT 900
 DB 841 GATGATTACGTCGCAAAAAGACTAAATGCTTAAAGCGGTGCGGAATATTTTGTGCT 900
 QY 901 GTTGAAGAGGTGAGACATCAGCTCTTGTGGGGTTTGAACCGGTGAAGATTATGAAT 960
 DB 901 GTTGAAGAGGTGAGACATCAGCTCTTGTGGGGTTTGAACCGGTGAAGATTATGAAT 960
 QY 961 GAGGCGTTTTTGAATGACAGCTGCTGACAGGTTGATCTTTTGTAGTACCCCAAGC 1020
 DB 961 GAGGCGTTTTTGAATGACAGCTGCTGACAGGTTGATCTTTTGTAGTACCCCAAGC 1020
 QY 1021 AATATTCAGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
 DB 1021 AATATTCAGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
 QY 1081 TTTATTTGTTAAGAGACACACTTTTACAGATGCTGATTAAGCAATTCAGCAACTTCAG 1140
 DB 1081 TTTATTTGTTAAGAGACACACTTTTACAGATGCTGATTAAGCAATTCAGCAACTTCAG 1140
 QY 1141 CAGGCTAAGGTATGAGTATGAGATTCCTGGAGTTGTATGATACAGGAATATGAG 1200
 DB 1141 CAGGCTAAGGTATGAGTATGAGATTCCTGGAGTTGTATGATACAGGAATATGAG 1200
 QY 1201 GAGATTAACCTTGGCTTGAAGAAAGGGGACTCTGACACTGCTTATAGGCAAAAGTTGATGA 1260
 DB 1201 GAGATTAACCTTGGCTTGAAGAAAGGGGACTCTGACACTGCTTATAGGCAAAAGTTGATGA 1260
 QY 1261 TGCCGATATGCTGCTGCTTGAAGACAGTGAAGATTCAATATAGGAATCAGCTATTTG 1320
 DB 1261 TGCCGATATGCTGCTGCTTGAAGACAGTGAAGATTCAATATAGGAATCAGCTATTTG 1320
 QY 1321 GAGTTGTTTGAAGAAATCAATGCTGATGACATGATGATGATGATGATGATGATGATGATGATGAT 1380
 DB 1321 GAGTTGTTTGAAGAAATCAATGCTGATGACATGATGATGATGATGATGATGATGATGATGATGAT 1380
 QY 1381 TTGTTGAAACCTGCTGCTGCAAGGGGTTGCTTCTAGGTTGACAGACACAAAGATTA 1440
 DB 1381 TTGTTGAAACCTGCTGCTGCAAGGGGTTGCTTCTAGGTTGACAGACACAAAGATTA 1440
 QY 1441 AATTTTAACTCGGGGACTACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
 DB 1441 AATTTTAACTCGGGGACTACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
 QY 1501 GAGGTATGCAAGGGTTCTCTTCTAGCTGCTGCAAGTGGCAAGATTTGAGCCGAG 1560
 DB 1501 GAGGTATGCAAGGGTTCTCTTCTAGCTGCTGCAAGTGGCAAGATTTGAGCCGAG 1560
 QY 1561 CATGTGAAGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
 DB 1561 CATGTGAAGACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
 QY 1621 AGAACTCGGCTGAACCCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
 DB 1621 AGAACTCGGCTGAACCCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680
 QY 1681 AACAATGATGAGCCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
 DB 1681 AACAATGATGAGCCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
 QY 1741 GAAAACCTTGAAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800
 DB 1741 GAAAACCTTGAAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800
 QY 1801 GAACTACTGTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
 DB 1801 GAACTACTGTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
 QY 1861 GCTGCTGATGAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1920
 DB 1861 GCTGCTGATGAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1920

QY 1921 AGCTCATCTTTTCAACGACGAGATATGGTTTCTTCTATGAGATCTGATGTGCTACCATTA 1980
DB 1921 AGCTCATCTTTTCAACGACGAGATATGGTTTCTTCTATGAGATCTGATGTGCTACCATTA 1980
QY 1981 GGGTCAGTCAGAGCTGAGAGATTTGAGAACACTTCCCAAGATGATGCTAGACATGCGAGAG 2040
DB 1981 GGGTCAGTCAGAGCTGAGAGATTTGAGAACACTTCCCAAGATGATGCTAGAGACTGCGAGAG 2040
QY 2041 AATATAGATCCAGAGTGGCAGAGATTTAAGTCTCTGCTTTTGGGCTGATCACCGCATTA 2100
DB 2041 AATATAGATCCAGAGTGGCAGAGATTTAAGTCTCTGCTTTTGGGCTGATCACCGCATTA 2100
QY 2101 GAAATGTTACCAAGCTTTTGGATGGCGCAATGCTGAAAGATTGAGTCTGACAGACGACT 2160
DB 2101 GAAATGTTACCAAGCTTTTGGATGGCGCAATGCTGAAAGATTGAGTCTGACAGACGACT 2160
QY 2161 GAACTGGGCGAGCTTGGTGGTGTATGATATACACTGTTGAAACATATCTGTTGACAGT 2220
DB 2161 GAACTGGGCGAGCTTGGTGGTGTATGATATACACTGTTGAAACATATCTGTTGACAGT 2220
QY 2221 GTGACAGTCTCAGAGATGAGAACCCGTCCTGCTGAGAGCACTCTGAGAGAGTCTGCT 2280
DB 2221 GTGACAGTCTCAGAGATGAGAACCCGTCCTGCTGAGAGCACTCTGAGAGAGTCTGCT 2280
QY 2281 TGCTTATCTGATTTGGTTCATCCAGAAAACAATGCTACTGATGTGCAAACTTACACAACA 2340
DB 2281 TGCTTATCTGATTTGGTTCATCCAGAAAACAATGCTACTGATGTGCAAACTTACACAACA 2340
QY 2341 AGATACGAGAGTTTCTGGTCCAGTCAGGGTGGAAAAATCACTGAAAGCTCTGTTGCA 2400
DB 2341 AGATACGAGAGTTTCTGGTCCAGTCAGGGTGGAAAAATCACTGAAAGCTCTGTTGCA 2400
QY 2401 TCATTA 2406
DB 2401 TCATTA 2406

RESULT 4
ADJ38210
ID ADJ38210 standard; DNA; 2637 BP.
XX
AC ADJ38210;
XX
DT 06-MAY-2004 (first entry)
XX
DE Plastid division-related Arc6 orthologue gene 3.
XX
KW prokaryotic type; plastid division; Ftn2; ARC6; ARC5; Fzo; plant cell;
KM agronomic; horticultural; crop plant; ornamental plant; woody plant;
KW herbicide target; gene; ds.
XX
OS Arabidopsis thaliana.
XX
PN MO2004001003-A2.
XX
PD 31-DEC-2003.
XX
PF 20-JUN-2003; 2003WO-US019536.
XX
PR 20-JUN-2002; 2002US-0390140P.
PR 09-AUG-2002; 2002US-0402242P.
PR 20-JUN-2003; 2003US-00600070.
XX
PA (UNMS) UNIV MICHIGAN STATE.
XX
PI Oseeryoung KW, Vitha S, Kokeharova OA, Gao H;
XX
DR WPI; 2004-082486/08.
XX
PT P-PSDB; ADJ38209.
XX
PT New isolated Ftn2, ARCS and/or Fzo-like nucleic acid sequences, useful
PT for further characterizing plastid division in plant cells, and in
PT varying agronomic and horticultural characteristics of economically

PT important plants.
XX
PS Disclosure; Fig 8; 287pp; English.
XX
CC This invention relates to novel prokaryotic type or plastid division and
CC related genes and proteins. In particular, the invention relates to novel
CC Ftn2 (ARC6), ARCS and Fzo-like genes and polypeptides. The methods and
CC compositions of the present invention are useful for further
CC characterizing plastid division in plant cells, in order to vary
CC agronomic and horticultural characteristics of economically important
CC plants, such as crop, ornamental and woody plants. They can also be used
CC as herbicide targets. The present sequence is that of a gene which is
CC related to the invention.
XX
SQ Sequence 2637 BP; 706 A; 535 C; 644 G; 752 T; 0 U; 0 Other;

Query Match 99.7%; Score 2399.6; DB 12; Length 2637;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2402; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGGAAGCTCTGAGTCAGTCGCGCATTTGGTCTCTCCCATTTCCAAATTATGCCATTACCA 60
DB 114 ATGGAAGCTCTGAGTCAGTCGCGCATTTGGTCTCTCCCATTTCCAAATTATGCCATTACCA 173
QY 61 CCGGCGAGCAAAAGCTCGACGCTGACCAACAACCTCTACAACTATCTGCTCCGACG 120
DB 174 CCGGCGAGCAAAAGCTCGACGCTGACCAACAACCTCTACAACTATCTGCTCCGACG 233
QY 121 AATATGGGCGACCGTCTCTCTCCAGTTCATTTACACCTCGATTCCTCTCCCTCC 180
DB 234 AATATGGGCGACCGTCTCTCTCCAGTTCATTTACACCTCGATTCCTCTCCCTCC 293
QY 181 TTGCGCACCGCCACCAACACCGCACCTCTGCTCTCTGACACCATATTTGATCGTCCC 240
DB 294 TTGCGCACCGCCACCAACACCGCACCTCTGCTCTCTGACACCATATTTGATCGTCCC 353
QY 241 GAACGCCAGCTCCCATCCCATTTGATTTTACACAGATTTAGAGAGCTCAAAACATTTTC 300
DB 354 GAACGCCAGCTCCCATCCCATTTGATTTTACACAGATTTAGAGAGCTCAAAACATTTTC 413
QY 301 TTAAACGATGAGATCCAGAAAGACATTCGAGGTTTCGAAACCGCGCAATTCGT 360
DB 414 TTAAACGATGAGATCCAGAAAGACATTCGAGGTTTCGAAACCGCGCAATTCGT 473
QY 361 TTGACGACGACGCTTTAATCAACCGGAGACAGATTTCTTCAAGCTGCTTGCAGAACTCTG 420
DB 474 TTGACGACGACGCTTTAATCAACCGGAGACAGATTTCTTCAAGCTGCTTGCAGAACTCTG 533
QY 421 TCTAATCTCGGCTAGAGAGATACATGAAGGTTCTTGTATGATGAAGAGCTTACA 480
DB 534 TCTAATCTCGGCTAGAGAGATACATGAAGGTTCTTGTATGATGAAGAGCTTACA 593
QY 481 GTGATCATGATGTCCTTGGGATTAAGGTTCCGGGCTCTGCTGATATGCAAGAGGT 540
DB 594 GTGATCATGATGTCCTTGGGATTAAGGTTCCGGGCTCTGCTGATATGCAAGAGGT 653
QY 541 GGTGAGACTGAGATGTTCTTCCGGTGTGAGAGCTTGTCTTAAGAGAGGTTGCCCTAG 600
DB 654 GGTGAGACTGAGATGTTCTTCCGGTGTGAGAGCTTGTCTTAAGAGAGGTTGCCCTAG 713
QY 601 TCGTTTAAGCAAGATGTGTTTATGATTAAGCGCTTGTCTCGATGTCTGAGAGGAT 660
DB 714 TCGTTTAAGCAAGATGTGTTTATGATTAAGCGCTTGTCTCGATGTCTGAGAGGAT 773
QY 661 GCTATGGCATTTGATCCACCTGATTTTATGCTTATAGTTTGTGTAAGAGGCTTTG 720
DB 774 GCTATGGCATTTGATCCACCTGATTTTATGCTTATAGTTTGTGTAAGAGGCTTTG 833
QY 721 AAGCTTTTACAGAGAGAGAGCAAGTATGCTTGCACCGGATTTTACGTCGACAAATTGAT 780
DB 834 AAGCTTTTACAGAGAGAGAGCAAGTATGCTTGCACCGGATTTTACGTCGACAAATTGAT 893
QY 781 GAGACTTTGAGAGAGATCACTCCGCGTTATGTCTTGGAGCTACTTGCGCTTACCGCTTGGT 840

```

Db      894 GAGACTTTGGAAGAGATCACTCGCGTTATGTCCTTGAGACTACTGGCTTACCGCTTGGT 953
Qy      841 GATGATTTACGCTCGCAAAAAGACTAAATGGTTTAAAGCGGTGTGCGAAATATTTTGTGTC 900
Db      954 GATGATTTACGCTCGCAAAAAGACTAAATGGTTTAAAGCGGTGTGCGAAATATTTTGTGTC 1013
Qy      901 GTTGAAGAGGTGAGCATCAGCTCTTGTGTGGGGTTTGAACCGGTGAGAAATTTATGAAT 960
Db      1014 GTTGAAGAGGTGAGCATCAGCTCTTGTGTGGGGTTTGAACCGGTGAGAAATTTATGAAT 1073
Qy      961 GAGGGCTTTTTCGAAATGACAGTGTGAGAGGTGATCTTTTGTACTACCCCAAC 1020
Db      1074 GAGGGCTTTTTCGAAATGACAGTGTGAGAGGTGATCTTTTGTACTACCCCAAC 1133
Qy      1021 AATATTCAGAGAGTCACTTTGAAAGTTTACGAAGTTGACCTTGCTGTGCTCAAGCT 1080
Db      1134 AATATTCAGAGAGTCACTTTGAAAGTTTACGAAGTTGACCTTGCTGTGCTCAAGCT 1193
Qy      1081 TTTATTTGTAAGAACGACACCTTTTACAGATGCTGATAGCAATTCAGCAACTTCAG 1140
Db      1194 TTTATTTGTAAGAACGACACCTTTTACAGATGCTGATAGCAATTCAGCAACTTCAG 1253
Qy      1141 CAGGCTAAGGTATGCTATGAGATTCCTGCGATGTTGTATGATACAGCAATTAATGG 1200
Db      1254 CAGGCTAAGGTATGCTATGAGATTCCTGCGATGTTGTATGATACAGCAATTAATGG 1313
Qy      1201 GAGATAGACTTCGCTTAGAAAAGGGGACTGTGACAGCTTATATGGCAAGTTGATGA 1260
Db      1314 GAGATAGACTTCGCTTAGAAAAGGGGACTGTGACAGCTTATATGGCAAGTTGATGA 1373
Qy      1261 TGCCGTATGTTGGGCTTAGACAGTGAAGATTCACATATAGAAATCAGCTATTTGG 1320
Db      1374 TGCCGTATGTTGGGCTTAGACAGTGAAGATTCACATATAGAAATCAGCTATTTGG 1433
Qy      1321 GAGTTGTTTGGAAATTCAAATGTGATGCAATGATGATCTCCCTGCACTATGCAAA 1380
Db      1434 GAGTTGTTTGGAAATTCAAATGTGATGCAATGATGATCTCCCTGCACTATGCAAA 1493
Qy      1381 TTGTTGAAAACCTGGTGGCAGGGGTGTCCTTCCAGGTTGAGAGACCAAGATTA 1440
Db      1494 TTGTTGAAAACCTGGTGGCAGGGGTGTCCTTCCAGGTTGAGAGACCAAGATTA 1553
Qy      1441 AAATTTAACTCGGGGACTACTATGATGATCCATGTTTGAAGTACTTGGAAAAGTG 1500
Db      1554 AAATTTAACTCGGGGACTACTATGATGATCCATGTTTGAAGTACTTGGAAAAGTG 1613
Qy      1501 GAGGTAGTTCAAGGTTCTCTTACCTGCTGCTGCACTATGCGCAAGATTGAGCCGAG 1560
Db      1614 GAGGTAGTTCAAGGTTCTCTTACCTGCTGCTGCACTATGCGCAAGATTGAGCCGAG 1673
Qy      1561 CATGTGAAGCTAGTGTATGACAGGCACTGCAAAAAGTTTTCTTCCCGCTATACAGT 1620
Db      1674 CATGTGAAGCTAGTGTATGACAGGCACTGCAAAAAGTTTTCTTCCCGCTATACAGT 1733
Qy      1621 AGAAACTCGGCTGAACCCAGAGATGTGAAGACAGTGTATGATGATCTTGTGGT 1680
Db      1734 AGAAACTCGGCTGAACCCAGAGATGTGAAGACAGTGTATGATGATCTTGTGGT 1793
Qy      1681 AACATATGAGCCGCTGATGATGAGCTGTGTCCTTATTTGCAAAAGCTGTAAACCTCT 1740
Db      1794 AACATATGAGCCGCTGATGATGAGCTGTGTCCTTATTTGCAAAAGCTGTAAACCTCT 1853
Qy      1741 GAAAACCTTTGAAAACCTATGATTAATGCAATTCAGAGCTGGGGGTCTCAGAGAGTGAAGCTTAT 1800
Db      1854 GAAAACCTTTGAAAACCTATGATTAATGCAATTCAGAGCTGGGGGTCTCAGAGAGTGAAGCTTAT 1913
Qy      1801 GAAACTACTGTTGAAATGTCGCTGCTGATATGTTAAAGAGCAAGTGTGAAGATCTTA 1860
Db      1914 GAAACTACTGTTGAAATGTCGCTGCTGATATGTTAAAGAGCAAGTGTGAAGATCTTA 1973
Qy      1861 GCTGCTGTGTGGCAATGGAATGATTTCACTGTTCAAGCCAGAAATTTTCTTAAAGC 1920

```

```

Db      1974 GCTGCTGTGTGGCAATGGAATGGAATTTCACTGTTCAAGCCAGAAATTTTCTTAAAGC 2033
Qy      1921 AGCTCATCTTTTCAACGCAAGGATATGTTTCTTCTATGGAATGATGCTGCTACATA 1980
Db      2034 AGCTCATCTTTTCAACGCAAGGATATGTTTCTTCTATGGAATGATGCTGCTACATA 2093
Qy      1981 GGGTCAGTCAAGCTGACGATTCAGAAAGCACTTCCAGAAATGATGCTGAGACTGAGAG 2040
Db      2094 GGGTCAGTCAAGCTGACGATTCAGAAAGCACTTCCAGAAATGATGCTGAGACTGAGAG 2153
Qy      2041 AATATAGTATCCAAATGTCAGAAAGTAAATGTCCTGCTTTTGGGCTGTATACCCGATA 2100
Db      2154 AATATAGTATCCAAATGTCAGAAAGTAAATGTCCTGCTTTTGGGCTGTATACCCGATA 2213
Qy      2101 GAAATGTTACAGAGGTTTGTGATGGGCGAATGCTGAAGATTGGAAGTGAAGAGAGCT 2160
Db      2214 GAAATGTTACAGAGGTTTGTGATGGGCGAATGCTGAAGATTGGAAGTGAAGAGAGCT 2273
Qy      2161 GAACTGCGCAGCTTGGGTTTATGATTAATTAACACTGTTGAAAATATCTGTTGACAGT 2220
Db      2274 GAACTGCGCAGCTTGGGTTTATGATTAATTAACACTGTTGAAAATATCTGTTGACAGT 2333
Qy      2221 GTGACAGTCTCAGCAGATGGAACCCGCTGCTGTGTGGAAGCACTCTGAGAGATCTGCT 2280
Db      2334 GTGACAGTCTCAGCAGATGGAACCCGCTGCTGTGTGGAAGCACTCTGAGAGATCTGCT 2393
Qy      2281 TGTCTATCTGATTTGGTTCAATCAGAAAACAATGCTACTGATGTGAGAACCTACACA 2340
Db      2394 TGTCTATCTGATTTGGTTCAATCAGAAAACAATGCTACTGATGTGAGAACCTACACA 2453
Qy      2341 AGATACGAAGTTTCTGTGTCAGATCAGGGTGAAGAAATCACTGAAAGGCTCTGTTTGA 2400
Db      2454 AGATACGAAGTTTCTGTGTCAGATCAGGGTGAAGAAATCACTGAAAGGCTCTGTTTGA 2513
Qy      2401 TCATTA 2406
Db      2514 TCATTA 2519

```

```

RESULT 5
ADT14901
ID ADT14901 standard; cDNA, 2679 BP.
XX
AC ADT14901;
XX
DT 13-JAN-2005 (first entry)
XX
DE Plant cDNA, Seq ID 227.
XX
KW Plant; ss; gene; transgenic; cold tolerance; growth rate;
KW drought tolerance; disease resistance; galactonmanan production;
KW plant growth regulator; heat tolerance; herbicide tolerance;
KW lignin production; extreme osmotic condition tolerance;
KW pathogens resistance; pest resistance; yield improvement; seed oil yield;
KW seed protein yield.
XX
OS Viridiplantae.
XX
PN US2004216190-A1.
XX
PD 28-OCT-2004.
XX
PF 18-DEC-2003; 2003US-00739930.
XX
PR 28-APR-2003; 2003US-00424599.
XX
PI 28-APR-2003; 2003US-00425115.
XX
PA (KOVA/) KOVALIC D K.
XX
PI KOVALIC DK;
XX
DR WPI; 2004-757369/74.
XX

```

PT New recombinant DNA constructs useful in the field of biochemistry and
PT genetics, and in particular for producing transgenic plants with improved
PT biological characteristics.

XX Claim 1; SEQ ID NO 227; 14bp; English.

XX The invention relates a recombinant DNA construct comprising a
CC polynucleotide having any of 5544 nucleotide sequences (cDNAs SEQ ID NO:
CC 1-5544) and encoding a polypeptide with any of 5544 amino acid sequences
CC (SEQ ID NO: 5545-11088). The cDNAs and proteins are from corn, soybean,
CC Arabidopsis, wheat and rape but the specification does not indicate which
CC sequences is derived from which organism. Also included is a method of
CC producing a plant having an improved property, comprising transforming a
CC plant with a recombinant DNA construct comprising a promoter region
CC functional in a plant cell operably joined to a polynucleotide encoding a
CC polypeptide associated with the property, and growing the transformed
CC plant. The property is selected from improving plant cold tolerance, for
CC manipulating growth rate in plant cells by modification of the cell cycle
CC pathway, for improving plant drought tolerance, for providing increased
CC resistance to plant disease, for galactomannan production, for production
CC of plant growth regulators, for improving plant heat tolerance, for
CC improving plant tolerance to herbicides, for increasing the rate of
CC homologous recombination in plants, for lignin production, for improving
CC plant tolerance to extreme osmotic conditions, for improving plant
CC tolerance to pathogens or pests, for yield improvement by modification of
CC photosynthesis, for modifying seed oil yield and/or content, for
CC modifying seed protein yield and/or content, for yield improvement by
CC modification of carbohydrate, nitrogen or phosphorus use and/or uptake
CC and for yield improvement by providing improved plant growth and
CC development under at least one stress condition. The polynucleotide may
CC also encode a plant transcription factor. The methods and compositions of
CC the present invention are useful in the field of biochemistry and
CC genetics, in particular for producing transgenic plants with improved
CC biological characteristics such as increased yield, improved nitrogen
CC flow, increasing plant tolerance to cold or heat, improving plant
CC tolerance to extreme osmotic and drought conditions, and improving plant
CC tolerance to plant pests or pathogens. They can also be used in physical
CC arrays of molecules, plant breeding markers, computer-based storage and
CC analysis systems. The present sequence is one of the 5544 plant cDNA
CC sequences of the invention. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?docid=20040216190.

XX Sequence 2679 BP; 712 A; 545 C; 653 G; 769 T; 0 U; 0 Other;

Query Match 99.7%; Score 2399.6; DB 13; Length 2679;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 2402; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGGAAGCTTGAAGTCAAGTGGGATTGGTCTCTCCCATTCGAATTATGCGGATTACA 60
DB 115 ATGGAAGCTTGAAGTCAAGTGGGATTGGTCTCTCCCATTCGAATTATGCGGATTACA 174
QY 61 CCGGCGAGCAAGAGCTCCGACGTAAGCAACAACCTCTACACTATCTGTCGGCAGC 120
DB 175 CCGGCGAGCAAGAGCTCCGACGTAAGCAACAACCTCTACACTATCTGTCGGCAGC 234
QY 121 AAATGGGCGGACGCTTCTCTCGACTTCATTTACCTCCGATTCCTCTCTCTCC 180
DB 235 AAATGGGCGGACGCTTCTCTCGACTTCATTTACCTCCGATTCCTCTCTCTCTCC 294
QY 181 TTGCGCACCGGACCAACCGGACCTCTGCTCTCTGCGACCATATTTGATGATGCC 240
DB 295 TTGCGCACCGGACCAACCGGACCTCTGCTCTCTGCGACCATATTTGATGATGCC 354
QY 241 GAACGCGACGTCCTCCATTCCTCCATTTCTTACACAGTATTTAGAGCTCAAAACATTTT 300
DB 355 GAACGCGACGTCCTCCATTCCTCCATTTCTTACACAGTATTTAGAGCTCAAAACATTTT 414
QY 301 TTAACCGATGGAATCAGAAAGACATTCGAAGCTAGGGTTTGAACCGCGCAATTCGGT 360
DB 415 TTAACCGATGGAATCAGAAAGACATTCGAAGCTAGGGTTTGAACCGCGCAATTCGGT 474

QY 361 TTCAGCGACGACGCTTTATATCAGCCGAGACAGATTCTTCAAGCTGCTCGAAACTCTG 420
DB 475 TTCAGCGACGACGCTTTATATCAGCCGAGACAGATTCTTCAAGCTGCTCGAAACTCTG 534
QY 421 TCTAATCTCGGCTCTAAGAGAGTACATGAAGGCTTTCTTGATGATGAAGAGCTACA 480
DB 535 TCTAATCTCGGCTCTAAGAGAGTACATGAAGGCTTTCTTGATGATGAAGAGCTACA 594
QY 481 GTCATCATGATGTTCTTGAGATTAAGTCTCGAGGCTCTGCTGATTAAGCAAGAGGT 540
DB 595 GTCATCATGATGTTCTTGAGATTAAGTCTCGAGGCTCTGCTGATTAAGCAAGAGGT 654
QY 541 GGTGAGACTGAGATGATCTTCTGCGGTTGTGAGGCTCTGCTTAAGAGAGGTTGCCTAAG 600
DB 655 GGTGAGACTGAGATGATCTTCTGCGGTTGTGAGGCTCTGCTTAAGAGAGGTTGCCTAAG 714
QY 601 TCGTTTAAGCAAGATGCTGTTTAAGTAAAGCCCTGCTTCTCGATGCTCGAGGGAT 660
DB 715 TCGTTTAAGCAAGATGCTGTTTAAGTAAAGCCCTGCTTCTCGATGCTCGAGGGAT 774
QY 661 GCTATGACATTTGATCCACCTGATTTTATTAAGTATGATGATGTTGTTGAGAAAGCTTTG 720
DB 775 GCTATGACATTTGATCCACCTGATTTTATTAAGTATGATGATGTTGTTGAGAAAGCTTTG 834
QY 721 AAGCTTTTACAG 780
DB 835 AAGCTTTTACAG 894
QY 781 GAGACTTTGGAAGAGATCATCTCCGCGTATGCTTGAAGCTACTTGGCTTACCGCTTGT 840
DB 895 GAGACTTTGGAAGAGATCATCTCCGCGTATGCTTGAAGCTACTTGGCTTACCGCTTGT 954
QY 841 GATGATTAACCTCGCGAAGAAAGTAAATGTTTAAAGCGTGTGCGGAATATTTTGTGCT 900
DB 955 GATGATTAACCTCGCGAAGAAAGTAAATGTTTAAAGCGTGTGCGGAATATTTTGTGCT 1014
QY 901 GTTGGAGAGGTGAGAGATCAGCTCTTGTGCGGTTTGAACCCGTGGAAGATTATGAAAT 960
DB 1015 GTTGGAGAGGTGAGAGATCAGCTCTTGTGCGGTTTGAACCCGTGGAAGATTATGAAAT 1074
QY 961 GAGGCGTTTTTAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1020
DB 1075 GAGGCGTTTTTAAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1134
QY 1021 AATATTCACAGAGATCATTTGAAGTTTACGAAGTTGCACTTCTCTTGTGCTCAAGCT 1080
DB 1135 AATATTCACAGAGATCATTTGAAGTTTACGAAGTTGCACTTCTCTTGTGCTCAAGCT 1194
QY 1081 TTTATTTGTAAGAAAGCCACCTTTTACAGAGATGCTGATTAAGCAATTCAGCAACTTCAG 1140
DB 1195 TTTATTTGTAAGAAAGCCACCTTTTACAGAGATGCTGATTAAGCAATTCAGCAACTTCAG 1254
QY 1141 CAGGCTTAAGGTATGCTATGAGAGATCTCTGCGATGTTGATGATACACGAATTAATG 1200
DB 1255 CAGGCTTAAGGTATGCTATGAGAGATCTCTGCGATGTTGATGATACACGAATTAATG 1314
QY 1201 GAGATAGACTTCCGCTCTAGAAAGGAGACTGTGCACTGCTTATTAAGCAAGTTGATGAA 1260
DB 1315 GAGATAGACTTCCGCTCTAGAAAGGAGACTGTGCACTGCTTATTAAGCAAGTTGATGAA 1374
QY 1261 TGCCTGATGTTGTTGGGCTTTAGACAGTGAAGATTCACAAATATAGAAATCCAGTATTTG 1320
DB 1375 TGCCTGATGTTGTTGGGCTTTAGACAGTGAAGATTCACAAATATAGAAATCCAGTATTTG 1434
QY 1321 GAGTTTGTTTTGAAGATTCAAATCGTGAAGACAAATGATATCTCCCTGGAAGTATGAAA 1380
DB 1435 GAGTTTGTTTTGAAGATTCAAATCGTGAAGACAAATGATATCTCCCTGGAAGTATGAAA 1494
QY 1381 TTGTTGGAAGACTGTTGGAGAGGGTTGCTTCTTCAAGTTCAAGAGACCAAAAGATATA 1440
DB 1495 TTGTTGGAAGACTGTTGGAGAGGGTTGCTTCTTCAAGTTCAAGAGACCAAAAGATATA 1554

OY		1441	AAATTTAAACCTGGGAGCACTAATAATATCCTAAGTTTTGAGTACTGTGAAGAAGTG	1500
Db		1555	AATTTAAACTCGGGAGCACTATATATATTCCTATGTTTGAAGTACTTGGAAGAAGTG	1614
OY		1501	GAGGTAGTTCAGGGTTCCTTTAGCTGCTGTCGACCTATGCGAAGATTGGAGCCGAG	1560
Db		1615	GAGGTAGTTCAGGGTTCCTTTAGCTGCTGTCGACCTATGCGAAGATTGGAGCCGAG	1674
OY		1561	CATGTGAAGCTAAGTGCATATGCAAGGCACTGCAGAAAAGTTTTCTTCCCCTATACAGAT	1620
Db		1675	CATGTGAAGCTAAGTGCATATGCAAGGCACTGCAGAAAAGTTTTCTTCCCCTATACAGAT	1734
OY		1621	AGAAACTCGGCTGAACCAGATATGTCAMAGACAGTGTTAAGTAGATCCTGTGGT	1680
Db		1735	AGAAACTCGGCTGAACCAGATATGTCAMAGACAGTGTTAAGTAGATCCTGTGGT	1794
OY		1681	AACAATGAGGCCGTGATGTTGAGCCTGTGTCTTTATTGCAAGACGTATGAAGCCCTCT	1740
Db		1795	AACAATGAGGCCGTGATGTTGAGCCTGTGTCTTTATTGCAAGACGTATGAAGCCCTCT	1854
OY		1741	GAAAACTTTGAAACTTAATGATTATGCAATTGAGCTGGGGTCTCAGAGATGAGGTTGAT	1800
Db		1855	GAAAACTTTGAAACTTAATGATTATGCAATTGAGCTGGGGTCTCAGAGATGAGGTTGAT	1914
OY		1801	GAAACTACTGTGAAATATGTCGTTGCTCATATATGTTAAAGAGCGCAAGTGAATCCTA	1860
Db		1915	GAAACTACTGTGAAATATGTCGTTGCTCATATATGTTAAAGAGCGCAAGTGAATCCTA	1974
OY		1861	GCTGCTGGTGTGGCAATTGCACTGATTCACGTTCCAGCCAGAAATATTTTCTTAAAGC	1920
Db		1975	GCTGCTGGTGTGGCAATTGCACTGATTCACGTTCCAGCCAGAAATATTTTCTTAAAGC	2034
OY		1921	AGCTCATCTTTTCAAACGCAAGATATGTTTCTCTATGAAATCTGATATGCTAACATA	1980
Db		2035	AGCTCATCTTTTCAAACGCAAGATATGTTTCTCTATGAAATCTGATATGCTAACATA	2094
OY		1981	GGGTCACTCAGAGCTGACGATTCAGAAAGCACTTCCAGAAATGATGTCTAGCACTGACAG	2040
Db		2095	GGGTCACTCAGAGCTGACGATTCAGAAAGCACTTCCAGAAATGATGTCTAGCACTGACAG	2154
OY		2041	AATATATGATTCAAAGTGCAGAGATTBAATGATCTCGGCTTTTGGGCTGATCACCGATA	2100
Db		2155	AATATATGATTCAAAGTGCAGAGATTBAATGATCTCGGCTTTTGGGCTGATCACCGATA	2214
OY		2101	GAAATGTATCACAGAGTTTTGATGAGGGGAAATGCTGAAGATTGGAATGCAACAGACGCT	2160
Db		2215	GAAATGTATCACAGAGTTTTGATGAGGGGAAATGCTGAAGATTGGAATGCAACAGACGCT	2274
OY		2161	GAAACTGCGCAGCTTGGGTTGGTTTATGATATATACACTGTGAAACTATCTGTGACAGT	2220
Db		2275	GAAACTGCGCAGCTTGGGTTGGTTTATGATATATACACTGTGAAACTATCTGTGACAGT	2334
OY		2221	GTCACAGTCTCAGACAGATGAAACCCGTGCTCTGTGTGAAGCACTGTGAGAGAGCTTGCT	2280
Db		2335	GTCACAGTCTCAGACAGATGAAACCCGTGCTCTGTGTGAAGCACTGTGAGAGAGCTTGCT	2394
OY		2281	TGCTATCTGATTTGGTTCATCCAGAAAACAATGCTACTATGTCACAACTTACACAACA	2340
Db		2395	TGCTATCTGATTTGGTTCATCCAGAAAACAATGCTACTATGTCACAACTTACACAACA	2454
OY		2341	AGATACGAAGTTTTCTGTCCAACTCAGGGTGAAGAAATCACTGAAGGCTCTGTTCCTGCA	2400
Db		2455	AGATACGAAGTTTTCTGTCCAACTCAGGGTGAAGAAATCACTGAAGGCTCTGTTCCTGCA	2514
OY		2401	TCATPA 2406	
Db		2515	TCATPA 2520	

RESULT 6
ADJ38130
ID ADJ38130 standard; DNA; 3667 BP
XX

AC	ADJ38130;
XX	
DT	06-MAY-2004 (first entry)
XX	
DE	Arabidopsis thaliana Arc6-1 genomic DNA SegID2.
XX	
KW	prokaryotic type; plastid division; Ftn2; ARC6; ARC5; Fzo: plant cell;
KM	agronomic; horticultural; crop plant; ornamental plant; woody plant;
XX	herbicide target; db.
OS	Arabidopsis thaliana.
PV	MO2004001003-A2.
PD	31-DEC-2003.
XX	
PF	20-JUN-2003; 2003WO-US019536.
XX	
PR	20-JUN-2002; 2002US-0390140P.
PR	09-AUG-2002; 2002US-0402242P.
XX	
PR	20-JUN-2003; 2003US-00600070.
PA	(UNMS) UNIV MICHIGAN STATE.
XX	
EI	Oseeryoung KW, Vicha S, Kokharova OA, Gao H;
DR	WPI; 2004-082486/08.
DR	P-PSDB; ADJ38202.
XX	
PT	New isolated Ftn2, ARCS and/or Fzo-like nucleic acid sequences, useful
FT	for further characterizing plastid division in plant cells, and in
PT	varying agronomic and horticultural characteristics of economically
XX	important plants.
PS	Claim 1; SEQ ID NO 2; 287bp; English.
XX	
CC	This invention relates to novel prokaryotic type or plastid division and
CC	related genes and proteins. In particular, the invention relates to novel
CC	Ftn2 (AR65), ARCS and Fzo-like genes and polypeptides. The methods and
CC	compositions of the present invention are useful for further
CC	characterising plastid division in plant cells, in order to vary
CC	agronomic and horticultural characteristics of economically important
CC	plants, such as crop, ornamental and woody plants. They can also be used
CC	as herbicide targets. The present sequence is that of a DNA sequence
XX	which is related to the invention.
SQ	Sequence 3667 BF; 983 A; 670 C; 842 G; 1172 T; 0 U; 0 Other;
	Query Match 71.7%; Score 1724; DB 12; Length 3667;
	Best Local Similarity 79.2%; Pred. No. 0;
	Matches 2406; Conservative 0; Mismatches 0; Indels 632; Gaps 5;
QY	1 ATGGAAGCTTGAATCAGTCACGTGGCATTTGGTCTCTGCCATTCCAAATTATGCGCAATTACA 60
DB	481 ATGGAACCTTGAGTCAAGTCAGTCGGATTGGTCTCTGCCATTCCAAATTATGCGCAATTACA 540
QY	61 CCGGCAGCACAAAGCTCCGACGTAAGGCACAAACACTCTCAAATATCTGTCGGCAGC 120
DB	541 CCGGCAGCACAAAGCTCCGACGTAAGGCACAAACACTCTCAAATATCTGTCGGCAGC 600
QY	121 AAATGGGCGACGCGTCTTCTCTCGAATTCAATTTCACCTTCGATTCTCTCTCTCTCC 180
DB	601 AAATGGGCGACGCGTCTTCTCTCGAATTCAATTTCACCTTCGATTCTCTCTCTCTCC 660
QY	181 TTGGCCACCGCCACACACCACCGGACCTCTGTCTCTTGCCACCATTTATGATGCTGCC 240
DB	661 TTGGCCACCGCCACACACCACCGGACCTCTGTCTCTTGCCACCATTTATGATGCTGCC 720
QY	241 GAAGGCCACGTCCCGCATCCCATTGATTTTTCACAGGATTTAGAGGCTCAAAACATTTTC 300
DB	721 GAAGGCCACGTCCCGCATCCCATTGATTTTTCACAGGATTTAGAGGCTCAAAACATTTTC 780
QY	301 TTAACCGATGAATTCAGAAAGACATTCGAAGCTTAGGGTTTTGAAACCGCGCAATTGCGT 360

|||||
Db 781 TTAACCATGGAAATCAAGAGACATTCGAGCTGAGGTTTGAAACCGCCGCAATTCGGT 840
Qy 361 TTGAGGACAGACGCTTTAATCAGCCGAGACAGATTCTTCAAGCTGCTTGGAATCTG 420
Db 841 TTGAGGACAGACGCTTTAATCAGCCGAGACAGATTCTTCAAGCTGCTTGGAATCTG 900
Qy 421 TCTAATCCTCGGTCTAGAGAGAGATGACATGAAAGTCTTCTTGATGATGAGAACTACA 480
Db 901 TCTAATCCTCGGTCTAGAGAGAGATGACATGAAAGTCTTCTTGATGATGAGAACTACA 960
Qy 481 GTATCATCTGATGTTCTTGGAAT----- 504
Db 961 GTATCATCTGATGTTCTTGGAATGAGTAAATTCGATTCGGAATATAAAGTTTCTTC 1020
Qy 505 -----AAGTTCTGAGGCG 518
Db 1021 GTTTTAATTTTCATGAATGGATTAAGAGAGAACTTTTATCTAGTGAAAGTTCCTGAGGCG 1080
Qy 519 TCTCTGTGATATTGACAGAGAGTGGTGAGACTGAGATAGTTCTTCGGGTGGTGAGGCTCT 578
Db 1081 TCTCTGTGATATTGACAGAGAGTGGTGAGACTGAGATAGTTCTTCGGGTGGTGAGGCTCT 1140
Qy 579 GCTTAAAGAGAGTTCCTTAAGTCTTTAAGCAAGATGTGTTTTAGTTATGCGCTTTC 638
Db 1141 GCTTAAAGAGAGTTCCTTAAGTCTTTAAGCAAGATGTGTTTTAGTTATGCGCTTTC 1200
Qy 639 GTTTCGATGATCTCGAGGAGTGGTATGGCATGGATTCACCTGATTTTATTAATCTGCTTA 698
Db 1201 GTTTCGATGATCTCGAGGAGTGGTATGGCATGGATTCACCTGATTTTATTAATCTGCTTA 1260
Qy 699 TGAGTTTGTAGAGAGCTTTGAAGCTTTTAC----- 730
Db 1261 TGAGTTTGTAGAGAGCTTTGAAGCTTTTACAGGTAGTTGACTGTGCTTTGTAATTTG 1320
Qy 731 -----A 731
Db 1321 ACGAGCGTTGGCTTTATAGAACTTCTGATTTGATACTTTGTATTTGAGTCTTGTGTA 1380
Qy 732 GGAGAGAGAGAGCAAGTAGCTTTCGACCGGATTTACGTGCACAAATTTGAGACTTTTGA 791
Db 1381 GGAGAGAGAGAGCAAGTAGCTTTCGACCGGATTTACGTGCACAAATTTGAGACTTTTGA 1440
Qy 792 AGAGTCACTCCGCTTATGTCTTGAGACTAATTGGCTTACCGCTGGTGAATGAATGACG 851
Db 1441 AGAGTCACTCCGCTTATGTCTTGAGACTAATTGGCTTACCGCTGGTGAATGAATGACG 1500
Qy 852 TCGCAAAAGACTAATGTGTTAAGCGGTGCGGAATTTTGTGTCTGTGTGAGAGAG 911
Db 1501 TCGCAAAAGACTAATGTGTTAAGCGGTGCGGAATTTTGTGTCTGTGTGAGAGAG 1560
Qy 912 TGGAGCATCAGCTCTTGTGGGGGTTTGACCGGTGAGAGATTTGATGATGAGGCGTTTT 971
Db 1561 TGGAGCATCAGCTCTTGTGGGGGTTTGACCGGTGAGAGATTTGATGATGAGGCGTTTT 1620
Qy 972 ACGATGACAGCTGCTGAGC----- 991
Db 1621 ACGATGACAGCTGCTGAGCAGGTATACAGTTTATGATACCTTTTATTTTCTTTAGCA 1680
Qy 992 -----AGTTGATCTTTTGTGA 1008
Db 1681 TGATATACTTTAGGTTTCTCATTTTAAATGATGTTGTGTGTTGAGTGTGATCTTTTGTGA 1740
Qy 1009 GCTACCCCAAGCAATATTCAGAGAGTCAATTTGAGATTTGAGAGTTGACTGCTCTT 1068
Db 1741 GCTACCCCAAGCAATATTCAGAGAGTCAATTTGAGATTTGAGAGTTGACTGCTCTT 1800
Qy 1069 GTGCTCAAGCTTTTATTTGTTAGAGAGCACCTTTTACAGATGCTGATAGCAATTC 1128
Db 1801 GTGCTCAAGCTTTTATTTGTTAGAGAGCACCTTTTACAGATGCTGATAGCAATTC 1180
Qy 1129 CAGCACTTCAGAGGCTTAAGTAAATGCTATGAGAGATTCCTGCGATGTTGTATGATACA 1188
|||||

Db 1861 CAGCACTTCAGAGGCTTAAGTAAATGCTATGAGAGATTCCTGCGATGTTGTATGATACA 1920
Qy 1189 CGGAATTAATTGGAGAGATAGACTTCGCGTCTAGAAAGGGAATCTGTGCACTGTTATAGCG 1248
Db 1921 CGGAATTAATTGGAGAGATAGACTTCGCGTCTAGAAAGGGAATCTGTGCACTGTTATAGCG 1980
Qy 1249 AAAGTTGATGAATGCCGATATGTGTTGGGCTTAGAACAGTGAAGATTCACAAATATAGGAAT 1308
Db 1981 AAAGTTGATGAATGCCGATATGTGTTGGGCTTAGAACAGTGAAGATTCACAAATATAGGAAT 2040
Qy 1309 CCAAGCTAATGCGAGTTGTTTGGAGAAATTCAAATCGGATGACAAATGATGATCTCCCT 1368
Db 2041 CCAAGCTAATGCGAGTTGTTTGGAGAAATTCAAATCGGATGACAAATGATGATCTCCCT 2100
Qy 1369 GGAATATGCAAAATTTGTTGAGAACTCGGTTGGCAGGGGTTGTCTTCTAGATTACAGAGAC 1428
Db 2101 GGAATATGCAAAATTTGTTGAGAACTCGGTTGGCAGGGGTTGTCTTCTAGATTACAGAGAC 2160
Qy 1429 ACCAAGATTAATAATTTTAACTCGGGGACTACTATGATGATCCATGTTTGAATTAC 1488
Db 2161 ACCAAGATTAATAATTTTAACTCGGGGACTACTATGATGATCCATGTTTGAATTAC 2220
Qy 1489 TTGGAAGAAGTGAAGGATGATTCAGGGGTTCTCTTATGCTGCTGCTGCACTATGCGAAG 1548
Db 2221 TTGGAAGAAGTGAAGGATGATTCAGGGGTTCTCTTATGCTGCTGCTGCACTATGCGAAG 2280
Qy 1549 ATTGAGCCGAGCATGTAAGCTAGTCTATGCAAGCACTGAGAAAGTTTTCCTTCC 1608
Db 2281 ATTGAGCCGAGCATGTAAGCTAGTCTATGCAAGCACTGAGAAAGTTTTCCTTCC 2340
Qy 1609 CGCTATACAGATAGAAATCTCGCTGAAACCAAGATGTGCAAGAGACAGTGTATGTGA 1668
Db 2341 CGCTATACAGATAGAAATCTCGCTGAAACCAAGATGTGCAAGAGACAGTGTATGTGA 2400
Qy 1669 GATCCGTTGGTAAACAATGAGGCCGATGAGGCTGAGGCTGCTGCTTATTTGAGAGAGCT 1728
Db 2401 GATCCGTTGGTAAACAATGAGGCCGATGAGGCTGAGGCTGCTGCTTATTTGAGAGAGCT 2460
Qy 1729 GTAAGACCCCTCTGAAAACTTTGAAACTAATGATTAATGCAATTCGAGCTGGGCTCTCAGAG 1788
Db 2461 GTAAGACCCCTCTGAAAACTTTGAAACTAATGATTAATGCAATTCGAGCTGGGCTCTCAGAG 2520
Qy 1789 AGTAGCGTTGATGAATCTAGTTGAAATGTCCGTTGCTGATATGTTAAAGAGAGCAAGT 1848
Db 2521 AGTAGCGTTGATGAATCTAGTTGAAATGTCCGTTGCTGATATGTTAAAGAGAGCAAGT 2580
Qy 1849 GTGAAGATCCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1908
Db 2581 GTGAAGATCCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2640
Qy 1909 TTTCTTAAAGAGCTCATCTTTTCAACGCAAGATATGTTTTCTTATGATGATCTGAT 1968
Db 2641 TTTCTTAAAGAGCTCATCTTTTCAACGCAAGATATGTTTTCTTATGATGATCTGAT 2700
Qy 1969 GTGCGTACCA----- 1978
Db 2701 GTGCGTACCAATGAGTATGATTAATGATGCAATTTTCAATATCTGATGCTCAATAATA 2760
Qy 1979 ----- 1978
Db 2761 TGCTTTTGTGAGCTTAAGACATATGTTCCACTTAATACATGTCCTCCAAAGTTGTACC 2820
Qy 1979 ----- 1978
Db 2821 AAGATTACAAGTTTGTGATGAATATTTCACTAATATATGCTGTGAATTTTGTGATCAA 2880
Qy 1979 ----- 1978
Db 2881 CTGTAGACAGAAATGTAATTTCACTCTCAACATTTCTGTGTAGAAATGATAGATAG 2940
Qy 1979 -----T 1979
Db 2941 AGATTGCTTAATGATGCTTTGTGCAACTTTTCTTTCCTGATTTTCTTTTGCATTT 3000
|||||

QY	1980	AGGGTCAGTCAGACCTGCACATTCAGAAAGCACTCCCGAATGAGTGTAGACATCGACA	203.9
Dp	3001	AGGGTCAGTCAGACCTGCACATTCAGAAAGCACTCCCGAATGAGTGTAGACATCGACA	306.0
QY	2040	GAATATATGATCCAAATGCGCAGAAAGATTAACTCTCGCTTTTGGGCTGTATCAACCGCAT	209.99
Dp	3061	GAATATATGATCCAAATGCGCAGAAAGATTAACTCTCGCTTTTGGGCTGTATCAACCGCAT	312.0
QY	2100	AGAAATGTTACAG-----	211.3
Dp	3121	AGAAATGTTACAGAGGTGAGGGAATAATCTACAATTCAAATGTGTGAAAACCTGT	318.0
QY	2114	-----AGGTTTGGATGGC	212.8
Dp	3181	TGCAATGATTTATATAGTCGTGGCCCTGTTGATTCTGTATTATTAAGGTTTGGATGGC	324.0
QY	2129	GAATGCTGAAGATTTGGACCTGACAGAGAGCGTGAACCTGGGCACTGGGGTTGGTTATG	218.8
Dp	3241	GAATGCTGAAGATTTGGACCTGACAGAGAGCGTGAACCTGGGCACTGGGGTTGGTTATG	330.0
QY	2189	ATTATACACTGTTGAAACTATCTGTTTGAACAGTGTGACAGTCTTGACAGATGGAACCCGTG	224.8
Dp	3301	ATTATACACTGTTGAAACTATCTGTTTGAACAGTGTGACAGTCTTGACAGATGGAACCCGTG	336.0
QY	2249	CTCGGTGGGAAGCAACTGTGGAGGAGCTGTGTTGTCTATCTGATTTGGTTCAATCCAAATA	230.8
Dp	3361	CTCGGTGGGAAGCAACTGTGGAGGAGCTGTGTTGTCTATCTGATTTGGTTCAATCCAAATA	342.0
QY	2309	ACAATGCTACTGATGTGAGAACTTACACAACAACAGATACGAATTTCTGTGTCAAGTCAG	236.8
Dp	3421	ACAATGCTACTGATGTGAGAACTTACACAACAACAGATACGAATTTCTGTGTCAAGTCAG	348.0
QY	2369	GGTGGAAATCACTGAAGGCTCTGTTCTGTGATATATA	240.6
Dp	3481	GGTGGAAATCACTGAAGGCTCTGTTCTGTGATATATA	351.8

Accession	Species	Gene	Source	Reference
ADJ38136	standard	DNA	3667 BP.	
ADJ38136	Arabidopsis thaliana	AtFtn2 genomic DNA	SeqID10.	
ADJ38136	Arabidopsis thaliana	AtFtn2 genomic DNA	SeqID10.	
06-MAY-2004	(first entry)			
prokaryotic type; plastid division; Ftn2, ARC6, ARC5, Fzo; plant cell;				
agronomic; horticultural; crop plant; ornamental plant; woody plant;				
herbicide target; ds.				
Arabidopsis thaliana.				
MO2004001003-A2.				
31-DEC-2003.				
20-JUN-2003; 2003WO-US019536.				
20-JUN-2002; 2002US-0390140P.				
09-AUG-2002; 2002US-0402242P.				
20-JUN-2003; 2003US-00600070.				
(UNMS) UNIV MICHIGAN STATE.				
Osteryoung KW, Viltha S, Koksharova OA, Gao H;				
WPI; 2004-082486/08.				
P-PSDB; ADJ38203.				
New isolated Ftn2, ARC5 and/or Fzo-like nucleic acid sequences, useful				
for further characterizing plastid division in plant cells, and in				

PT Varying agronomic and horticultural characteristics of economically
PT important plants.
XX
PS Claim 1; SEQ ID NO 10; 287pp; English.

CC This invention relates to novel prokaryotic type or plasmid division and
CC related genes and proteins. In particular, the invention relates to novel
CC *PutI* (Arc61), *Arc5* and *Pzo*-like genes and polypeptides. The methods and
CC compositions of the present invention are useful for further
CC characterizing plasmid division in plant cells, in order to vary
CC agronomic and horticultural characteristics of economically important
CC plants such as crop, ornamental and woody plants. They can also be used
CC as herbicide targets. The present sequence is that of a DNA sequence
CC which is related to the invention.

Sequence 3667 BP; 982 A; 669 C; 843 G; 1173 T; 0 U; 0 Other;

Query Match	71.58;	Score 1720.8;	DB 12;	Length 3667;
-------------	--------	---------------	--------	--------------

Matches 2404; Conservative 0; Mismatches 2; Indels 632; Gaps 5;

QY 1 ATGGAAGCTCTGAGTCACGTCGGCAATTGGTCTCTCCCAATTCGAATTATGCCGATTACCA 60

Db 481 ATGGAAGCTCTGAGTCACGTCGGCAATTGCTCTCTCCCAATTCGAATTATGCCGATTACCA 540

61 CCGCGACGACAAAGCTCCGACGTAGCCACAACACCTTACAACTATCTGCTCCGCCAGC 120

Db 541 CCGGCGACGACAAAGCTCCGACGTAGCCACAACACTTCTGCTCCGCCAGC 600

QY 121 AAATGGGCGACCGTCTTCTCTCGACTTCAATTTACACCTCGGATTCCTCCTCCTCC 180

Db 601 AAATGGGCGACCGTCTTCTCTCCGACTTCAATTTACACCTCCGATTCCTCCTCCTCC 660

181 TTGCCACCGCCACCACTCTGCTCTCTGCCACCATCTATTGATCGTCCC 240

Db 661 TTGGCCACCGCCACCAACCGCACTCTGTCCTCTGCCACCATCTATGATCGTCCC 720

241 GAAAGCCACGTCCTCCCATCCCATTTGATTTCTACCGGTAATTAGGAGCTCAAAACACATTTTC 300

Db 721 GAACGCCACGTCCCCCATCCCCCATTTGATTCTACCGATATTAGGAGCTCAAAACACATTC 780

301 TTACCGATGGAATCAGAAGCATTCGAAGCTAGGGTTTCGAAACCGCCGCAATTCCGT 360

Db 781 TTAACCGATGGAATCAGAAGAGCATTCGAAGCTAGGGTTTCGAAACCGCCGCAATTCCGT 840

361 TTCAGCGACGACGCTTTAATCAGCCGGAGACAGATTCTTCAAGCTGCTTGCAGAACTCTG 420

Db 841 TTCAGCGACGCGCTTAAATCAGCCGGAGACAGATTCTTCAAGCTGCTTGGGAAACTCTG 900

421 TCTAATCCTCGGTCTAGAAGAGTACAATGAAGTCTTCTTGATGATGAAGAAGCTACA 480

Db 901 TCTAATCCTCGGCTCTAGAGAGAGTACCAATGAAAGGCTCTTCTTGATGATGAAAGAGCTACA 960

481 GTCATCACTGATGTTCCCTTGGAT----- 504

D6 961 GTCATCAGTGTTCCTTGGGATAAGTAATTTCGATTTCCGATAATAAAGTTCTTC 1020

QY 505 -----AAGGTTCCCTGGGC 518

D5 1021 GTTTAATT CATGAATGGAAGGA CTTTATCTAGTGAAGTTCCTGGGC 1080

519 TCTCTGTGATTTGCAGAGGCTGGTGAGACTGAGATAGTTCCTCCGGTTGGTGAGGCTCT 578

DB 1081 TCTCTGTAATGCCAAGAGGTGGTGGAGACTGAGATAGTCTCTCCGGTGGTGGAGGCTCT 1140

5/9 GCTTAAAGGAGGTTGCTTAAAGTCGTTTAAAGCAAGATGTGGTTTAAAGTATAGCTTATGGCCCTTGC 638

1141 GCTTAAAGGAGGTTGCTTAAGTCTTAAAGCAAGATGTGGTTTAACTTAAAGGCGCTTGC 1200

639 GTTCTCGAGTCTCGAGGATGCTATGGCAATGGATCCACCTGATTTATACGGTA 698

DB 1201 GTTCTCGATGTCGAGGATGCTATGGCAATGGATCCACCTGATTATTACTGGTAA 1260

DB 3421 ACAATGCTACTGATGTGACAGACCTTACACAAAGATACGAAGTTTCTGGTCCAAAGTAG 3480
QY 2369 GGTGGAATCACTGAGGCTCTGTTCTGCATCATTA 2406
DB 3481 GGTGGAATCACTGAGGCTCTGTTCTGCATCATTA 3518
RESULT 8
ADJ38212/c
ID ADJ38212 standard; cDNA; 561 BP.
XX ADJ38212;
AC
XX
XX 06-MAY-2004 (first entry)
DE Plastid division-related Arc6 orthologue cDNA 5.
XX
XX prokaryotic type; plastid division; Ftn2; ARC6; ARC5; Fzo; plant cell;
KM agronomic; horticultural; crop plant; ornamental plant; woody plant;
XX herbicide target; gene; ss.
OS Arabidopsis thaliana.
XX
XX WO2004001003-A2.
PN 31-DEC-2003.
XX
XX 20-JUN-2003; 2003WO-US019536.
PF
XX 20-JUN-2002; 2002US-0390140P.
PR 09-AUG-2002; 2002US-0402242P.
PR 20-JUN-2003; 2003US-00600070.
XX
XX (UNMS) UNIV MICHIGAN STATE.
XX
XX Osteryoung KW, Vittha S, Koksharova OA, Gao H;
XX WPI; 2004-082486/08.
DR
XX
XX New isolated Ftn2, ARC5 and/or Fzo-like nucleic acid sequences, useful
PT for further characterizing plastid division in plant cells, and in
PT varying agronomic and horticultural characteristics of economically
PT important plants.
XX
XX Disclosure; Fig 8; 287pp; English.
XX
XX
XX This invention relates to novel prokaryotic type or plastid division and
CC related genes and proteins. In particular, the invention relates to novel
CC Ftn2 (ARC6), ARC5 and Fzo-like genes and polypeptides. The methods and
CC compositions of the present invention are useful for further
CC characterizing plastid division in plant cells, in order to vary
CC agronomic and horticultural characteristics of economically important
CC plants, such as crop, ornamental and woody plants. They can also be used
CC as herbicide targets. The present sequence is a cDNA sequence which is
CC related to the invention.
XX
XX
SQ Sequence 561 BP; 167 A; 136 C; 107 G; 148 T; 0 U; 3 Other;
Query Match 20.0%; Score 482; DB 12; Length 561;
Best Local Similarity 99.4%; Pred. No. 1.5e-130;
Matches 482; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1922 GCTCATCTTTCAAGCAAGATAGTGTCTTCTATGGAATCTGATGCGTACCATAG 1981
DB 561 GCTCATCTTTCAAGCAAGATAGTGTCTTCTATGGAATCTGATGCGTACCATAG 502
QY 1992 GGTCACTGACGCTGACGATTCAGAGCACTTCCAGAAATGATGCTGAGACTGACAGA 2041
DB 501 GGTCACTGACGCTGACGATTCAGAGCACTTCCAGAAATGATGCTGAGACTGACAGA 442
QY 2042 ATATGTAATCCAAAGTGGAGAGATTAAGTCTCTGCTTTGGGCTGATCCCGCATAG 2101
DB 441 ATATGTAATCCAAAGTGGAGAGATTAAGTCTCTGCTTTGGGCTGATCCCGCATAG 382

QY 2102 AAATGTTACGAGAGTTTGGATGGGCGAATGCTGAAGATTGGAAGTGAAGAGAGCTG 2161
DB 381 AAATGTTACGAGAGTTTGGATGGGCGAATGCTGAAGATTGGAAGTGAAGAGAGCTG 322
QY 2162 AAATGTTACGAGAGTTTGGATGGGCGAATGCTGAAGATTGGAAGTGAAGAGAGCTG 2221
DB 321 AAATGTTACGAGAGTTTGGATGGGCGAATGCTGAAGATTGGAAGTGAAGAGAGCTG 262
QY 2222 TGACAGTCTGACGAGATGGAACCGGTCTCTGATGGAAGCACTGTGAGAGAGTGTGCTT 2281
DB 261 TGACAGTCTGACGAGATGGAACCGGTCTCTGATGGAAGCACTGTGAGAGAGTGTGCTT 202
QY 2282 GCTATCTGATTTGGTTTCATCAGAAAACAAATGCTATGATGTCAGAACTTACACAA 2341
DB 201 GCTATCTGATTTGGTTTCATCAGAAAACAAATGCTATGATGTCAGAACTTACACAA 142
QY 2342 GATACGAAGTTTTCGTTGTCGCAAGTCAGAGTGAAGAAATCACTGAAGGCTCTGTTGCAT 2401
DB 141 GATACGAAGTTTTCGTTGTCGCAAGTCAGAGTGAAGAAATCACTGAAGGCTCTGTTGCAT 82
QY 2402 CATAA 2406
DB 81 CATAA 77
RESULT 9
ADJ38206
ID ADJ38206 standard; DNA; 2283 BP.
XX
XX ADJ38206;
AC
XX
XX 06-MAY-2004 (first entry)
DT
XX
XX Plastid division-related Arc6 orthologue gene 1.
DE
XX
XX prokaryotic type; plastid division; Ftn2; ARC6; ARC5; Fzo; plant cell;
KM agronomic; horticultural; crop plant; ornamental plant; woody plant;
XX herbicide target; gene; ds.
OS Oryza sativa.
XX
XX WO2004001003-A2.
FN
XX
XX 31-DEC-2003.
PD
XX 20-JUN-2003; 2003WO-US019536.
PF
XX 20-JUN-2002; 2002US-0390140P.
PR 09-AUG-2002; 2002US-0402242P.
PR 20-JUN-2003; 2003US-00600070.
XX
XX (UNMS) UNIV MICHIGAN STATE.
XX
XX Osteryoung KW, Vittha S, Koksharova OA, Gao H;
XX WPI; 2004-082486/08.
DR P-PSDB; ADJ38205.
XX
XX New isolated Ftn2, ARC5 and/or Fzo-like nucleic acid sequences, useful
PT for further characterizing plastid division in plant cells, and in
PT varying agronomic and horticultural characteristics of economically
PT important plants.
XX
XX Disclosure; Fig 8; 287pp; English.
XX
XX
XX This invention relates to novel prokaryotic type or plastid division and
CC related genes and proteins. In particular, the invention relates to novel
CC Ftn2 (ARC6), ARC5 and Fzo-like genes and polypeptides. The methods and
CC compositions of the present invention are useful for further
CC characterizing plastid division in plant cells, in order to vary
CC agronomic and horticultural characteristics of economically important
CC plants, such as crop, ornamental and woody plants. They can also be used

CC as herbicide targets. The present sequence is that of a gene which is
CC related to the invention.

XX Sequence 2283 BP; 551 A; 576 C; 592 G; 564 T; 0 U; 0 Other;

Query Match 20.0%; Score 481; DB 12; Length 2283;

Best Local Similarity 54.3%; Pred. No. 6.7e-130;

Matches 1256; Conservative 0; Mismatches 880; Indels 177; Gaps 7;

```
Qy 108 CTGCTCCGCGCAAGATGGCGGACCGCTTCTCTCCGACATTTCATTTCCACTCCGATTC 167
Db |||||
Qy 132 CTGGGCCAAGCGCTCTTGCGCGACTTCCACTCTCCACCGCGCCCTCCGACC 191
Db |||||
Qy 168 CTCTCTCTCTCTCTCTGCGCACCGCACCAACCGCACTCTGCTCTCTGCAACATC 227
Db |||||
Qy 192 GCGGTCCCGGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTT 251
Db |||||
Qy 228 TATTGATGTCGCCGAAGCGCACGTCCTCCATCCCATTTGATTTCTACAGATTAAGAGC 287
Db |||||
Qy 252 CCCCCAGCCCGCGAAGCGCTCTCCGCTCCAGTCGATTTCTACAGGTTCTAGGGGC 311
Db |||||
Qy 288 TCACAACATTTCTTAACGATGGAATGAGAAAGCATTCGAAAGCTAGGGTTGAAAC 347
Db |||||
Qy 312 AGAGCACATTTCTTGCGAGATGATCAAGAGGCGTTGAGGCAAGATTCGCAAGCC 371
Db |||||
Qy 348 GCGGCAATTCGGTTTCAAGCGACGACGCTTTAATCAGCCGAGACAGATTTCTCAAGCTGC 407
Db |||||
Qy 372 ACCGAGATGCGCTACAGACAGGATGCTCTTGTTGTCGTGCAAAATGCTGCAAGATTGC 431
Db |||||
Qy 408 TTGCGAAACTCTGTATTCCTCGGCTAGAAAGAGATCAATGAAGTCTTTCTTGATGA 467
Db |||||
Qy 432 CCATGACCTCTCATGAACCAAGAACTCCGCACTCAGATGATGCTGCGCTTTCTGAGAA 491
Db |||||
Qy 468 TGAAGAAGCTACAGTCACTGATGTCCTTGGAATAGGTCCTGGGGCTCTCTGATG 527
Db |||||
Qy 492 CCGTGAAGAGCTCTCACCATGATATTTGTTGSAAGAGAGGCTGGG----- 540
Db |||||
Qy 528 ATTGCAAGAGGTGCTGAGACTGAGATAGTTCTTTCGGGTTGGTGAAGCTTGCTTAAGA 587
Db |||||
Qy 541 -----GAGGCACTTGCTGTGTTGTTAATCTGAGAAAGAGTTGCTTCTGGA 584
Db |||||
Qy 588 GAGGTGCTTAAGTGGTTTAAGCAAGATGAGTTTAACTGATAGCGCTTGGCTTTCTGA 647
Db |||||
Qy 585 TCGGCAACCAAGCGCTTCAAGCAAGAGAGTGTGCTACGATGCTCTGCTTAATGATGA 644
Db |||||
Qy 648 TGTCTCGAGGAGTGTATGSCATTTGATCCACCTGATTTTATTAATGCTGTTAAGTTGT 707
Db |||||
Qy 645 TCTATCAAGGAGTGTATGSCAGCAAGCCCTCCAGATGTAATGCTGCTCTCGAGATGCT 704
Db |||||
Qy 708 TGAGGAAGCTTTGAAGCTTTTACAGAGGAAGAGCAAGTACCTTGACCGGATTTACG 767
Db |||||
Qy 705 CGAGAGGCTCTCAAGGCTCTTGCAAGGAAGATGAGCAAGCAATCTCGACCTGATCTGCT 764
Db |||||
Qy 768 TGCACAAATTGATGAGACTTTGGAAGAGATCACTCCGCTTAATGCTTTGAGACTTGG 827
Db |||||
Qy 765 TTCAAGATTAATGAACCTCTGAGAGATTAACCTGCTGTGATTAATGAGCTTCTCTC 824
Db |||||
Qy 828 CTACCGCTGATGATTAACGCTGCGAAGAAAGTAAATGATTAAGCGGATGCGGAA 887
Db |||||
Qy 825 CTTTCTTATTAACAGAGCATCATTAAGAGCGCAAGAGGCTTCAAGGTTGCGAAGAA 884
Db |||||
Qy 888 TATTTTGTGCTGTGTTGAGAGAGTGAAGATCACTCTGTTGAGGCTTTGACCGCTGA 947
Db |||||
Qy 885 CATTTTGTGAGCGTTGCGAGAGGATTAATGCTACCGTTGAGAGAGATTTTCTCGTGA 944
Db |||||
Qy 948 GAAATTATGATGAGCGCTTTTACGATGACAGCTCTGAGAGGTTGATCTTTTGT 1007
Db |||||
Qy 945 AGCTTTCATGAACGAGCTTTTGTGAGGATGACATTAATGAACGATGAGATTTCTTTTC 1004
Db |||||
Qy 1008 AGCTACCCCAAGCAATATTCAGAGAGATCAATTTGAAGTTTACGAAGTTGACTTGGCT 1067
Db |||||
Qy 1005 AAAAAACCAATGACATTTCTCTGAAATGTTGAAAAATTAACATGATGACATTCGACAA 1064
Db |||||
```

```
Qy 1068 TGTGGCTCAAGCTTTTATTTGATAGAACCAACCTTTTACAGAGCTGATTAAGCAATT 1127
Db |||||
Qy 1065 TGTGCTCAAGCATTTATTAAGTAAAGGCCAATTCATCATGATGCGGATGATCTTTT 1124
Db |||||
Qy 1128 CCAGCAACTTCAGAGCTTAAGTAAATGCTATGAGAGATTCCTGCATGTTGATGATAC 1187
Db |||||
Qy 1125 TGAACCACTCCAGAAAGTTCAA-----CATTAAGTTCTCAATTA 1160
Db |||||
Qy 1188 ACGGATTAATTTGGAGATAGACTTCGCTAGAAAGGGAGCTGTCGACTGCTTAATAG 1247
Db |||||
Qy 1161 TGCTTATGATTAATGATGATGACCTTGATGAAAGGCAATTCGCTCATTTGCTAGTCGG 1220
Db |||||
Qy 1248 CAAAGTATGATGATGCGGATGATGATGAGCTTAAAGACATGAGAGATTCACAAATAGAA 1307
Db |||||
Qy 1221 AGATGTTAGCAAGTACGAATGATGCTTGAATTTGATTAATGAGTCTTCAACATACAGAA 1280
Db |||||
Qy 1308 TCCAGCTATTTGTGAGATTTGTTTGGAGATTCAAATC---GTGATCAATATGATGCT 1364
Db |||||
Qy 1281 CCCCCAAATTTCTAGAGTTTATGTGACCAACTGAGCATGAGTGAAGAGATGATCTTCT 1340
Db |||||
Qy 1365 CCTGGACTATGCAAAATTTGTTGAAAACCTGATTTGGCAGGGGTTGCTTTCTAGATTGAG 1424
Db |||||
Qy 1341 TCCAGGCTGTGCAAGCTTTTGGAGACTTGGCTTATCTTTGAGGTTTTTCTTAGAGACAG 1400
Db |||||
Qy 1425 AGACACCAAAAGATTAATAATTTAACTCGGAGCACTATGATGATCTATGCTTTTGAAG 1484
Db |||||
Qy 1401 AGATACCTGGGGGCAATGCAATTCAGACTTGAAGATTAATCAATGATGATCAGAAATTAAG 1460
Db |||||
Qy 1485 TTACTTGGAAAAGATGAGATGATTTCAAGGTTTCTCTTAACTGCTGCTGCAACTAATGC 1544
Db |||||
Qy 1461 CTACTTGAAGAAAGATGAGGATGATGCTTCAATTTGCTGCTGCTGCTGCTATATGC 1520
Db |||||
Qy 1545 AAGATTTGAGACCCGAGCATGTGAAGAGCTAGTCTATGCAAGCACTGCAAGAAATTTTCC 1604
Db |||||
Qy 1521 AAAACTTGTGTCTCAACTACACTGACGTGCACTTGGTA----- 1555
Db |||||
Qy 1605 TTCCCGCTATACAGATTAAGAACTCGGCTGAACCCAGAGATGTCAGAGACAAGTATTAG 1664
Db |||||
Qy 1556 -----CTGTAATATCAAAATGATTAATCAAGC----- 1581
Db |||||
Qy 1665 TGTATATCTGTTGTATTAACATGATGAGCCGTGATGATGAGCTGCTGCTTATATGAGA 1724
Db |||||
Qy 1582 -----TTCAACAA 1589
Db |||||
Qy 1725 AGCTTAAGACCCCTCTGAAAACCTTGAACCTAATGATTAATGCAATTCAGCTGGGGCTGC 1784
Db |||||
Qy 1590 GCTTTTCCATGATTAAGAACATGTTAGCAGCTGACCATGAAAAATCTAAGATGCGCC 1649
Db |||||
Qy 1785 AGAGATGAGCGTTGATGAACCTACTGTGAAATGTCGGTGTGATGATGTTAAAGAGGC 1844
Db |||||
Qy 1650 TGGGGATATCTTGAAATTTTGAACAGAAATGCACTGCTCATGATTCGAGAAATGTC 1709
Db |||||
Qy 1845 AAGTGTAAATCTTACTGCTGCTGTGSCAATTTGACATGATTTCACTGTTCAAGCCGAA 1904
Db |||||
Qy 1710 CGCTTAAGATTAATCTCTGTGCGCACTGTTTGAACATTTGACATTAATGAGGCGCAA 1769
Db |||||
Qy 1905 GTATTTCTTAAAAGCACTCATCTTTTCAACGCAAGAGATAGGTTCTTCTATGATAC 1964
Db |||||
Qy 1770 ATATTTGCTCTGTAAGAGGCGCCCTTCTGCTATTAAGAGTGAAGATGATCTGTGGAGT 1829
Db |||||
Qy 1965 TGAT-----GTGCTACATAGGGTCAATGATGAGCTGACGATTCAGAGCACTTCC 2015
Db |||||
Qy 2016 CAGAAATGATGCTAGACTGACAGAAATTAATGATTAATGATGATGATGATGATGATGAT 2075
Db |||||
Qy 1990 TAGATGATGATGAGGCTGAGAGATGATGATGATGATGATGATGATGATGATGATGAT 1949
Db |||||
Qy 2076 GGGTTTTGGGCTGATACCGCATAGAAATGTTTACAGAGTTTGGATGGGCGAATGCT 2135
Db |||||
Qy 1950 GGCCTTGGACCAAGAACTTCGTTGATCATTTGCAAGAGATTTCTTGAATGCAACATGCT 2009
Db |||||
Qy 2136 GAAATTTGACATGACAGACAGCTGAACCTGCGCAGCTTGGGTTGTTATGATTAATAC 2195
Db |||||
```

```
Db 2010 AAGGCTGAGCTGACCGAGCAGCGAGATTGAGCGCTCATTGGGTGTTCTGGAGATATC 2069
Qy 2196 ACTGTTGAACATATCTGTTGACAGTGTGACAGTGTGACAGATGAAACCCGTTCTGT 2255
Db 2070 ACTATCCAGATGTGACGATTGATGATCATCTATCTCCATGATGTGCGAGCCAGACTGT 2129
Qy 2256 GGAAGCACTGTGAGAGCTGTCTTCTATCTGATTTGTTTCATCCAGAAAACATGTC 2315
Db 2130 GGAGGCTACGATTGATGAGGAGCCAGCACTTACTGATGTACTGAGCCAGAAAACATGA 2189
Qy 2316 TACTGATGTGACAGCTTACACAAAGATGCAAGTTTCTGTCCAG---TCAGGGTG 2372
Db 2190 TTCATATGACACAAAATACATACCGGTATGATGAGTGGCTTCTCCAGCTAGAGGGTG 2249
Qy 2373 GAAATCACTGAAAGCTGTCTTCTGTCATCATTA 2405
Db 2250 GAGATACGGAAGAGCAGTCTCTCAAGTCGTA 2282

RESULT 10
ADJ38264
ID ADJ38264 standard; cDNA; 631 BP.
XX
AC ADJ38264;
XX
DT 06-MAY-2004 (first entry)
XX
DE Plastid division-related Arc6 orthologue cDNA 51.
XX
KM prokaryotic type; plastid division; Ftn2; ARC6; ARC5; Fzo; plant cell;
KW agronomic; horticultural; crop plant; ornamental plant; woody plant;
KW herbicide target; gene; ss.
XX
OS Prunus persica.
XX
PN WO2004001003-A2.
XX
PD 31-DEC-2003.
XX
PF 20-JUN-2003; 2003WO-US019536.
XX
PR 20-JUN-2002; 2002US-0390140P.
PR 09-AUG-2002; 2002US-0402242P.
PR 20-JUN-2003; 2003US-0060070.
XX
PA (UNMS ) UNIV MICHIGAN STATE.
XX
PI Osteeryoung KW, Vitha S, Koksharova OA, Gao H;
XX
DR WPI; 2004-082486/08.
XX
PT New isolated Ftn2, ARC5 and/or Fzo-like nucleic acid sequences, useful
PT for further characterizing plastid division in plant cells, and in
PT varying agronomic and horticultural characteristics of economically
PT important plants.
XX
PS Disclosure; Fig 8; 287pp; English.
XX
XX This invention relates to novel prokaryotic type or plastid division and
XX related genes and proteins. In particular, the invention relates to novel
XX Ftn2 (ARC6), ARC5 and Fzo-like genes and polypeptides. The methods and
XX compositions of the present invention are useful for further
XX characterizing plastid division in plant cells, in order to vary
XX agronomic and horticultural characteristics of economically important
XX plants, such as crop, ornamental and woody plants. They can also be used
XX as herbicide targets. The present sequence is a cDNA sequence which is
XX related to the invention.
XX
SQ Sequence 631 BP; 174 A; 123 C; 155 G; 178 T; 0 U; 1 Other;
Query Match 13.7%; Score 328.8; DB 12; Length 631;
Best Local Similarity 71.5%; Pred. No. 1.8e-85;
```

```
Matches 451; Conservative 0; Mismatches 168; Indels 12; Gaps 1;
Qy 913 GGAGCATGAGCTCTTGTGGGGGTTTGAACCCGTGAGAGTTTATGATGAGGCGTTTGA 972
Db 1 GCAAGTTGCAATTTCTGGGGGAGATTCTACTGTGAAAATTTTCAATGAAGAGGCTTTCTTG 60
Qy 973 CGAATGACAGCTGCTGAGCAGGTTGATCTTTTGTAGCTTACCCCAAGCAATATTCACGA 1032
Db 61 CATATGACTGAGCTGAGCAGGTTGATTTATTTGAGTACCCCAAGTAATATCCCGCA 120
Qy 1033 GAGTCATTTGAAGTTTAAAGTTGCACTTCTGTGGGCTGCAAGCTTTTATTTGGTAAG 1092
Db 121 GAAAGCTTTGAAGTTTATGAGGAGTGGCTTGTGCGCTGTGTGCTCAAGCTTTGTGGTAA 180
Qy 1093 AAGCCACACTTTTACAGAGATGCTGATTAAGCAATTCAGCAACTTCAGCAGGCTTAAGSTA 1152
Db 181 AAACCTCATCATTTCAAGATGCTGAAAACCTATTTCCAGAAAACCTTCAGCAGCTTAAGSTA 240
Qy 1153 ATGGCTATGAGAATTCCTGCGATGTTGTATGATACACGGAATTAATTGGAGATAGACTTC 1212
Db 241 ACAGCTGTAGGACATTTCTCTTGACAACTATATTAACCAAGAAAGCAGTGAATAGACTTT 300
Qy 1213 GGTCTAGAAAGGGGACTCTGTGCACTGTTTATAGCAAAAGTTGATGATGCCGTATGTGG 1272
Db 301 GCTTTGAGAGGGGACTCTGTTCATCTTCTAGGGGACTTGTATGACAGTCGTTCCGTGG 360
Qy 1273 TTGGGCTTGAACAGTGAAGATTCACAATATATGAATCCAGCTAATGTGGAGTTGTTTGG 1332
Db 361 TTGGGCTTGAACAGTGAATGATTCACTATATGAAATCATCTGTGTGACTTTGTCTTGG 420
Qy 1333 GAGAATTCAAATCGTGATG-----ACAATGATGATCTCCCTGACATATGCAAA 1380
Db 421 GAGAACTCAAGAGATGACATGATGACAAATGACAAATGATCTTCCGTGACCTTGGCAAG 480
Qy 1381 TTGTTGAAACCTGTTGGGAGGCTTCTTCTTCTAGTTTCAGAGACCAAAAGATATAA 1440
Db 481 CTATTTGAGAGCTGTGTATGATGAGGTTGATTTCCCAAGTTTAAAGACCAAAAGACATA 540
Qy 1441 AAATTTAACTCGGGGACTATCATGATGATCTCTATGTTTGAAGTTACTTGGAAAAGAGTG 1500
Db 541 GAGTTCAAGCTGGAGAGACTTATGATGATCTTCAAGCTTGAAGTACTTGAAGAAAGGCTG 600
Qy 1501 GAGGTAGTTCAAGGTTCTCTTCTTCTTCTGCTG 1531
Db 601 GATGGCACTAATGTTGACCTTGAAGCTGCTG 631

RESULT 11
ADJ38215
ID ADJ38215 standard; cDNA; 660 BP.
XX
AC ADJ38215;
XX
DT 06-MAY-2004 (first entry)
XX
DE Plastid division-related Arc6 orthologue cDNA 8.
XX
XX prokaryotic type; plastid division; Ftn2; ARC6; ARC5; Fzo; plant cell;
KW agronomic; horticultural; crop plant; ornamental plant; woody plant;
KW herbicide target; gene; ss.
XX
OS Medicago truncatula.
XX
PN WO2004001003-A2.
XX
PD 31-DEC-2003.
XX
PF 20-JUN-2003; 2003WO-US019536.
XX
PR 20-JUN-2002; 2002US-0390140P.
PR 09-AUG-2002; 2002US-0402242P.
PR 20-JUN-2003; 2003US-0060070.
XX
```

PA (UNMS) UNIV MICHIGAN STATE.
XX
XX Oosteryoung KW, Vitha S, Koksharova OA, Gao H;
XX MPI; 2004-082486/08.
DR
XX
XX
XX New isolated Ftn2, ARCS and/or Fzo-like nucleic acid sequences, useful
XX for further characterizing plastid division in plant cells, and in
XX varying agronomic and horticultural characteristics of economically
XX important plants.
XX
XX Disclosure: Fig 8; 287pp; English.
XX
XX This invention relates to novel prokaryotic type or plastid division and
XX related genes and proteins. In particular, the invention relates to novel
XX Ftn2 (ARCS), ARCS and Fzo-like genes and polypeptides. The methods and
XX compositions of the present invention are useful for further
XX characterizing plastid division in plant cells, in order to vary
XX agronomic and horticultural characteristics of economically important
XX plants, such as crop, ornamental and woody plants. They can also be used
XX as herbicide targets. The present sequence is a cDNA sequence which is
XX related to the invention.
XX
SQ Sequence 660 BP; 149 A; 178 C; 152 G; 181 T; 0 U; 0 Other;
Query Match 11.6%; Score 278.8; DB 12; Length 660;
Best Local Similarity 67.3%; Pred. No. 9.6e-71;
Matches 432; Conservative 0; Mismatches 192; Indels 18; Gaps 2;
QY 81 ACGTAGCCCAACACCTTACACATATGCTCCGCCAGCAATGGGCCGACCTTCT 140
DB 17 ACCTAACCGTCTCAATTCCTCCGCGTCCGCCACCGTAATGGGGGAGCACTCAT 76
QY 141 CTCGACTTCATTTACCTCCGATTCCTCTCCCTCCCTTCGCGCACCGCACACAC 200
DB 77 TTCGATTTCCATTTCCCTGGGAGACCTCTCTCTCTCCCTCCACACACTCCGAC 136
QY 201 CGCAGCTCTGCTCTCTGCCACCATCTAATGATGTCGCCAGCGCCAGTCCCATCC 260
DB 137 AGTCACTCTC-----ACTCTCTTACCTCCCTCCGATAGAGCGCAGTCACTCC 190
QY 261 CATTGATTTTACCGGATATAGAGGTCACCAACATTTCTTACCGATGGAATCAGAG 320
DB 191 TCTGACCTGTACCAAAATCTCGGCGCGCAACGATTTCTCGGTGGATTTCCGAG 250
QY 321 AGCATTCGAACTAGGGTTTGGAAACCGCCGCAATTCGGTTTCAGCAGCAGCTTAA 380
DB 251 AGCTTATGAAGCAATTCCTGAGGCTCTCAGATGCTTTCAATATGAGAGCTTAT 310
QY 381 CAGCCGAGACAGATTCTTCAAGCTGCTTGGAAACCTGTCTAATCTCGGCTAGAG 440
DB 311 TAGTCGTGTGAGATTCTTCAAGCTGCTTGGAAACCTAGCTGATCTCGCTTAGAG 370
QY 441 AGAGTACAAATGAAGTCTTCTTGATGATGAA-----GAGTACAGTCATCAG 488
DB 371 AGAGTATATGAAGCTCTGTGAGATGAAGACGAGATGAGATCTTCAATCTTCA 430
QY 489 TGATGTCCTTGGGATAGGTTCTCGGGGCTCTGTGTATTTGCAAGAGGTGGAGAC 548
DB 431 TGAATTCCTTTGCAAAAGTTCTCGAGCTCTGTGCTGTGTCAGAGAGCTGAGAGAC 490
QY 549 TGAATAGTTCTTGGGTTGTGAGGCTCTGCTTAAAGAGAGGTTCCCTAAGCTTAA 608
DB 491 GGAAGTGGTGTGCGGATTTGAGAGGGGTTTACTGAGAGAGAGTTTACGAGAGATGTTAA 550
QY 609 GCAAGATGTGTTTATGTTATGAGGCTGTCTGATGATCTCGAGAGATGCTATAGC 668
DB 551 GCAAGATGTGTGTGCTATGCGCTTGCAATGTTGACGTTCTTGAAGATGCTATAGC 610
QY 669 ATTGATCCACCTGATTTTATTAATGATGATGATGATGATGATGATGATGATGAT 710
DB 611 TTTGTCCCGGCAAGATTTTCAATTTGTTGTTGATGATGATGATGATGATGATGAT 652

RESULT 12
ADJ38223
ID ADJ38223 standard; cDNA; 537 BP.
XX
XX
XX ADJ38223;
XX
XX 06-MAY-2004 (first entry)
XX
XX
XX Plastid division-related Arc6 orthologue cDNA 16.
XX
XX
XX prokaryotic type; plastid division; Ftn2; ARCS; ARCS; Fzo; plant cell;
XX agronomic; horticultural; crop plant; ornamental plant; woody plant;
XX herbicide target; gene; ss.
XX
XX Triticum aestivum.
XX
XX MO2004001003-A2.
XX
XX
XX 31-DEC-2003.
XX
XX 20-JUN-2003; 2003WO-US019536.
XX
XX 20-JUN-2002; 2002US-0390140P.
XX 09-AUG-2002; 2002US-0402242P.
XX 20-JUN-2003; 2003US-00600070.
XX
XX (UNMS) UNIV MICHIGAN STATE.
XX
XX Oosteryoung KW, Vitha S, Koksharova OA, Gao H;
XX MPI; 2004-082486/08.
XX
XX
XX New isolated Ftn2, ARCS and/or Fzo-like nucleic acid sequences, useful
XX for further characterizing plastid division in plant cells, and in
XX varying agronomic and horticultural characteristics of economically
XX important plants.
XX
XX Disclosure: Fig 8; 287pp; English.
XX
XX
XX This invention relates to novel prokaryotic type or plastid division and
XX related genes and proteins. In particular, the invention relates to novel
XX Ftn2 (ARCS), ARCS and Fzo-like genes and polypeptides. The methods and
XX compositions of the present invention are useful for further
XX characterizing plastid division in plant cells, in order to vary
XX agronomic and horticultural characteristics of economically important
XX plants, such as crop, ornamental and woody plants. They can also be used
XX as herbicide targets. The present sequence is a cDNA sequence which is
XX related to the invention.
XX
SQ Sequence 537 BP; 133 A; 113 C; 151 G; 140 T; 0 U; 0 Other;
Query Match 9.3%; Score 224; DB 12; Length 537;
Best Local Similarity 63.6%; Pred. No. 1.1e-54;
Matches 341; Conservative 0; Mismatches 195; Indels 0; Gaps 0;
QY 555 AGTTCCTGGGTTGTGAGGCTCTGCTTATGAGAGAGGTTGATGCTTAAAGCAGA 614
DB 2 AGTTCCTGCAATTTGAGGGGCACTTACGAGAGACCGCCGCAAGGCTTCAAGCAGA 61
QY 615 TGTGTTTATGATAGGCGCTTGGTTTCTGATGTCTGAGGAGATGATGCAATTTGA 674
DB 62 TGTGTTGCTGGCAATGCGCTCTTATGTGATCTATCAAGGAGAGCAATGCGGCTAG 121
QY 675 TCAACCTGATTTTATTAATGCTTATGATGATGATGATGATGATGATGATGATGAT 734
DB 122 CCTCCAGATGATATCCGCTCTGTGTGAGTGTGAAAGGCTCTCAAGCTTTTGACGA 181
QY 735 GGAAGAGCAAGTAGCTTCCACCGATTTACGTGACCAATTTGATGAGATTTGAGAGA 794
DB 182 GGAATGGGCAATGATCTGACCTGGTTTGTCTGCAAAATTTGATGATGATGATGAT 241
QY 795 GATCACTCCGCGTTATGTCTTGAAGCTACTTGGCTTACCGCTTGGTATGATTAACGCTGC 854

Db 242 TATCACACCTCGTTGTTGTTGAGACCTTCTCCCTCTCTGATGAAAAACATCAGAA 301
QY 855 GAAAAGACAAATGTTTAAGCGGTGCGGAAATTTTGTGTCGTTGGAGAGTGG 914
Db 302 TGAACACCAAGAGGTCTTCGTGTGTGAAACATTTTGTGAGTGTGGACAGAGG 361
QY 915 AGCATCAGCTCTTGTGTTGGGGTTTGAACCCGTGAGAAAGTTTAAATGAGCCGTTTACG 974
Db 362 TATGAGTACTGTTGAGAGGAGATTTCCGTGAGGCTACATGAATGAAGCTTCTGTA 421
QY 975 AATGACAGTGTGAGCAGGTTGATCTTTTGTAGTACCCAGACATATTCGAGCAGA 1034
Db 422 GATGACATCGCGGAGCAGATGATTTCTTCAAAAACACCGAATAGCATCCGCTGA 481
QY 1035 GTCATTTGAAGTTTGAAGTTGACCTGCTGTGCGTCAAGCTTTTATTTGTA 1090
Db 482 ATGTTTGAATCTTATAGCGTGGCACTTGCAAAATGTGCTCAAGCAATTTGAATGA 537

RESULT 13

ACN48855/C

ID ACN48855 standard; cDNA; 552 BP.

AC ACN48855;

DT 02-DEC-2004 (first entry)

DE Cotton primed seed EST Clone ID: LIB3825-027-Q6-N6-H1, SEQ:3636.

XX Cotton; plant; EST; expressed sequence tag; transgenic plant; seed;
KM variety DP50B; library LIB3825; molecular tag; molecular marker;
KM genetic mapping; molecular mapping; seed germination; plant growth;
KM plant quality; plant yield; plant breeding; tissue printing; ss.

OS Gossypium hirsutum.

PN US2004123340-A1.

PD 24-JUN-2004.

PF 12-DEC-2001; 2001US-00021323.

PR 14-DEC-2000; 2000US-0255619P.

PA (DEIK/) DEIKMAN J.

PA (FENG/) FENG P C C.

PA (FINC/) FINCHER K L.

PA (ZIEG/) ZIEGLER T E.

PI Delkman J, Feng PCC, Fincher KL, Ziegler TE;

WPI; 2004-479808/45.

XX New isolated nucleic acid molecule that encodes a plant protein or its
PT fragment, useful for isolating a variety of agronomically significant
PT genes associated with plant growth, quality or yield, and as molecular
PT tags to map genes.

PS Claim 1; SEQ ID NO 3636; 34pp; English.

XX The invention relates to 17880 cotton expressed sequence tags (ESTs;
CC ACN45220-ACN63099). The ESTs were isolated from cDNA libraries generated
CC from primed or non-primed seeds from variety DP50B, mature seeds from
CC variety Coker 312 Boswell 96 Field, and androecium tissue, gynoecium
CC tissue, developing fibres, carpel walls and sepal from variety
CC Nucleon338. The invention also relates to substantially purified
CC proteins or their fragments encoded by nucleic acid molecules of the
CC invention, and to transformed plants having a nucleic acid construct
CC comprising a nucleic acid of the invention. The cotton ESTs are useful as
CC molecular tags to isolate genetic regions, to isolate genes, to map
CC genes, to determine gene function and to determining whether genes are
CC members of a particular gene family. The nucleic acid molecules may be

CC used for isolating a variety of agronomically significant genes
CC associated with plant growth, quality, yield, and could also serve as
CC links in metabolic and catabolic pathways. The nucleic acid molecules are
CC also useful for identifying genes important in initiating and maintaining
CC seed germination or that may be used to mitigate stresses encountered
CC during seed germination. The ESTs additionally enable the acquisition of
CC promoters and cis-regulatory elements which will be useful to express
CC agronomically significant genes in these tissues and/or other tissues,
CC and also permits the acquisition of molecular markers useful in breeding
CC schemes, genetic and molecular mapping, and in cloning of agronomically
CC significant genes. The nucleic acid molecules are further useful for
CC detecting the expression level or pattern of a protein or mRNA and for
CC detecting the presence or quantity of a protein by tissue printing. The
CC present sequence represents a specifically claimed EST isolated from a
CC cotton variety DP50B primed seed cDNA library (LIB3825). The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from the US patent office at
CC seqdata.uspto.gov/sequence.html?docID=US20040123340
XX

SQ Sequence 552 BP; 153 A; 122 C; 110 G; 167 T; 0 U; 0 Other;

Query Match 9.2%; Score 222.4; DB 13; Length 552;

Best Local Similarity 72.2%; Pred. No. 3.4e-54;

Matches 289; Conservative 0; Mismatches 111; Indels 0; Gaps 0;

QY 2006 AAGCACTTCCAGATGAGTCTAGAGCTGCGAGAAATATATATCCAAAGTGGCAGAGA 2065
Db 550 AAGATTAATCTAGATGATGATGCAAGAAATGCGAAGGCAATGTTGCGAAGTGGCAGATA 491

QY 2066 TTAAGTCTCTGCTTTTGGGCTGATACCGCATRAGAAATGTACAGAGTGTGGATG 2125
Db 490 TTAATATCGAGCGCTTTGAGCTTATACCCCTTGATTAATATGCGAAGGTTCTGAGTG 431

QY 2126 GCGCAATGCTGAAGTTTGAAGTGAAGAGGAGCTGAAACTGCGAGCTTGGTGGTGT 2185
Db 430 GTCAATATTTAAGACATGAGCAGATGTCGACGCCGAATGCTCAGCTTGGTGGTAT 371

QY 2186 ATGATTTATACCTGTGAACTATCTGTGACAGTGTGACAGTCTCAGACAGATGAGACC 2245
Db 370 ATGATATATAGCTACTGAAACATGCGCATGTACAGTGTACCTTTCACTAATGAGCCAGC 311

QY 2246 GTGCTGTGTTGAAAGCACTTGGAGAGTGTGCTGTCTATCTGATTTGGTTCATCCAG 2305
Db 310 GAGCTGTAGTGAAGCTACTCTGGAAGATCCACTGTGCTTACATGATTTATCATCCGG 251

QY 2306 AAAACAATGCTACTGATGTGCAACCTTACACACAAGATACAGTTTCTGTCCAGAT 2365
Db 250 AGAACAAATGCTCTAATGTAACTCTACACACAGATATGAGATGTCTTGTCCAACT 191

QY 2366 CAGGTTGGAATCACTGAAAGCTCTGTTCTTGATCATTA 2405
Db 190 CAGGCTGGAATATCACTGAAGATCTGTCTCAATCTTA 151

QY 2405 CAGGTTGGAATCACTGAAAGCTCTGTTCTTGATCATTA 2405
Db 151 CAGGCTGGAATATCACTGAAGATCTGTCTCAATCTTA 151

QY 2405 CAGGTTGGAATCACTGAAAGCTCTGTTCTTGATCATTA 2405
Db 151 CAGGCTGGAATATCACTGAAGATCTGTCTCAATCTTA 151

QY 2405 CAGGTTGGAATCACTGAAAGCTCTGTTCTTGATCATTA 2405
Db 151 CAGGCTGGAATATCACTGAAGATCTGTCTCAATCTTA 151

QY 2405 CAGGTTGGAATCACTGAAAGCTCTGTTCTTGATCATTA 2405
Db 151 CAGGCTGGAATATCACTGAAGATCTGTCTCAATCTTA 151

QY 2405 CAGGTTGGAATCACTGAAAGCTCTGTTCTTGATCATTA 2405
Db 151 CAGGCTGGAATATCACTGAAGATCTGTCTCAATCTTA 151

QY 2405 CAGGTTGGAATCACTGAAAGCTCTGTTCTTGATCATTA 2405
Db 151 CAGGCTGGAATATCACTGAAGATCTGTCTCAATCTTA 151

QY 2405 CAGGTTGGAATCACTGAAAGCTCTGTTCTTGATCATTA 2405
Db 151 CAGGCTGGAATATCACTGAAGATCTGTCTCAATCTTA 151

QY 2405 CAGGTTGGAATCACTGAAAGCTCTGTTCTTGATCATTA 2405
Db 151 CAGGCTGGAATATCACTGAAGATCTGTCTCAATCTTA 151

QY 2405 CAGGTTGGAATCACTGAAAGCTCTGTTCTTGATCATTA 2405
Db 151 CAGGCTGGAATATCACTGAAGATCTGTCTCAATCTTA 151

QY 2405 CAGGTTGGAATCACTGAAAGCTCTGTTCTTGATCATTA 2405
Db 151 CAGGCTGGAATATCACTGAAGATCTGTCTCAATCTTA 151

QY 2405 CAGGTTGGAATCACTGAAAGCTCTGTTCTTGATCATTA 2405
Db 151 CAGGCTGGAATATCACTGAAGATCTGTCTCAATCTTA 151

QY 2405 CAGGTTGGAATCACTGAAAGCTCTGTTCTTGATCATTA 2405
Db 151 CAGGCTGGAATATCACTGAAGATCTGTCTCAATCTTA 151

QY 2405 CAGGTTGGAATCACTGAAAGCTCTGTTCTTGATCATTA 2405
Db 151 CAGGCTGGAATATCACTGAAGATCTGTCTCAATCTTA 151

QY 2405 CAGGTTGGAATCACTGAAAGCTCTGTTCTTGATCATTA 2405
Db 151 CAGGCTGGAATATCACTGAAGATCTGTCTCAATCTTA 151

QY 2405 CAGGTTGGAATCACTGAAAGCTCTGTTCTTGATCATTA 2405
Db 151 CAGGCTGGAATATCACTGAAGATCTGTCTCAATCTTA 151

QY 2405 CAGGTTGGAATCACTGAAAGCTCTGTTCTTGATCATTA 2405
Db 151 CAGGCTGGAATATCACTGAAGATCTGTCTCAATCTTA 151

QY 2405 CAGGTTGGAATCACTGAAAGCTCTGTTCTTGATCATTA 2405
Db 151 CAGGCTGGAATATCACTGAAGATCTGTCTCAATCTTA 151

QY 2405 CAGGTTGGAATCACTGAAAGCTCTGTTCTTGATCATTA 2405
Db 151 CAGGCTGGAATATCACTGAAGATCTGTCTCAATCTTA 151

```
XX 20-JUN-2003; 2003WO-US019536.
PF
XX
PR 20-JUN-2002; 2002US-0390140P.
PR 09-AUG-2002; 2002US-0402242P.
PR 20-JUN-2003; 2003US-00600070.
XX
XX (UNMS ) UNIV MICHIGAN STATE.
PI Oosteryoung KW, Vittha S, Koksharova OA, Gao H;
XX WPI; 2004-082486/08.
XX
XX New isolated Ftn2, ARCS and/or Fzo-like nucleic acid sequences, useful
PT for further characterizing plastid division in plant cells, and in
PT varying agronomic and horticultural characteristics of economically
PT important plants.
XX
XX Disclosure; Fig 8; 287bp; English.
XX
XX This invention relates to novel prokaryotic type or plastid division and
CC related genes and proteins. In particular, the invention relates to novel
CC Ftn2 (ARCS), ARCS and Fzo-like genes and polypeptides. The methods and
CC compositions of the present invention are useful for further
CC characterising plastid division in plant cells, in order to vary
CC agronomic and horticultural characteristics of economically important
CC plants, such as crop, ornamental and woody plants. They can also be used
CC as herbicide targets. The present sequence is a cDNA sequence which is
CC related to the invention.
XX
XX Sequence 545 BP; 164 A; 107 C; 120 G; 151 T; 0 U; 3 Other;
SQ
Query Match 8.6%; Score 207; DB 12; Length 545;
Best Local Similarity 72.0%; Pred. No. 1.1e-49;
Matches 270; Conservative 0; Mismatches 105; Indels 0; Gaps 0;
OY 2032 ACTGAGAGATATATGATCCAAAGTGGCAAGATTAAAGTCTTGGCTTTGGCTTAT 2091
Db 2 ATTGACAGAGGCAATGTTGCGCAAGTGGCAGAACATTAAATCTGAGCGCTTGGACCTGAT 61
OY 2092 CACCGCATAGAAATGTTCACAGAGTTTGGATGGGCAATGCGAAGATTGGACCTGAC 2151
Db 62 CACCGCTTGAATTAATTCGCGAGAGTCTGGATGTCMAATGTTGAAGACATGGACAGAT 121
OY 2152 AGAGCAGGTGAAACTGCGCAGCTTGGTGTGTTATGATTAACACTGTTGAACCTATCT 2211
Db 122 CGTGAGCGCAAAATGCTCAGCTTGGGTATGATATAGTCTACTGAACATGGCC 181
OY 2212 GTTGACATGTGACAGTCTCAGACATGGAACCCGCTGCTGTGGTGAAGCAACTTGGAG 2271
Db 182 ATTGACAGTGTAACTCTTCACTAGATGCGCAGCAGCTGTAGTGAAGCTACTCTGGA 241
OY 2272 GAGTGTGCTTGTCTATGATTTGGTTCATCCAGAAAACAATGCTAGTGTGCAAGC 2331
Db 242 GAATCACTGCTTACTGATGATTTATATATCCGGAACAATGCTCTAATGTAATCTCC 301
OY 2332 TACACAAAGATACGAAGTTTCTGTGTCGAAGTCAAGGTGGAATAATCACTGAAGCTCT 2391
Db 302 TACACCAAGATATAGATGCTTGTTCAACTCAGGCTGGAATAATCACTGAAGATCT 361
OY 2392 GTTCTTGATCATTA 2406
Db 362 GTCTAACAATCTTAA 376
XX
XX RESULT 15
XX ID ADJ38254 standard; cDNA; 491 BP.
XX AC ADJ38254;
XX DT 06-MAY-2004 (first entry)
XX
```

```
DE Plastid division-related Arc6 orthologue cDNA 41.
XX
XX prokaryotic type; plastid division; Ftn2; ARCS; ARCS; Fzo; plant cell;
XX agronomic; horticultural; crop plant; ornamental plant; woody plant;
XX herbicide target; gene; ss.
XX
XX Triticum aestivum.
XX
XX WO2004001003-A2.
XX
XX 31-DEC-2003.
XX
XX 20-JUN-2003; 2003WO-US019536.
PF
XX
XX 20-JUN-2002; 2002US-0390140P.
PR 09-AUG-2002; 2002US-0402242P.
PR 20-JUN-2003; 2003US-00600070.
XX
XX (UNMS ) UNIV MICHIGAN STATE.
PA
XX Oosteryoung KW, Vittha S, Koksharova OA, Gao H;
XX WPI; 2004-082486/08.
XX
XX New isolated Ftn2, ARCS and/or Fzo-like nucleic acid sequences, useful
PT for further characterizing plastid division in plant cells, and in
PT varying agronomic and horticultural characteristics of economically
PT important plants.
XX
XX Disclosure; Fig 8; 287bp; English.
XX
XX This invention relates to novel prokaryotic type or plastid division and
CC related genes and proteins. In particular, the invention relates to novel
CC Ftn2 (ARCS), ARCS and Fzo-like genes and polypeptides. The methods and
CC compositions of the present invention are useful for further
CC characterising plastid division in plant cells, in order to vary
CC agronomic and horticultural characteristics of economically important
CC plants, such as crop, ornamental and woody plants. They can also be used
CC as herbicide targets. The present sequence is a cDNA sequence which is
CC related to the invention.
XX
XX Sequence 491 BP; 107 A; 123 C; 142 G; 111 T; 0 U; 8 Other;
SQ
Query Match 8.2%; Score 197.2; DB 12; Length 491;
Best Local Similarity 64.1%; Pred. No. 8.1e-47;
Matches 295; Conservative 0; Mismatches 165; Indels 0; Gaps 0;
OY 383 GCCGGAACAGATTTCTTCAAGCTGTTGCGAAACTCTGTCTAATCTCGGTCTGAAGAG 442
Db 2 GCCGTGGCAAAATCTGCAANTTGCACATGATATCTTCACAAACAGAGCTCCGCAACG 61
OY 443 AGTCAATGAAAGCTTCTTGATGATGAAGAAGCTACATCACTAGATTTCTTGGG 502
Db 62 AGTATGACCGCGCTCTGTGAGAACCGTACCGGGGCTCACACTGATGTTCTTGGG 121
OY 503 ATAAAGTTCTGGGGCTCTCTGTGTATTTGCAAGAGGTGTGAGACTGAGATAGTTCTTC 562
Db 122 ACAAGTTCCGGGGTGTCTATGTCCTTCAGAGAGCTGGGGAGGCAAGGAGTGTG 181
OY 563 GGGTTGTGAGGCTCTGCTTAAGAGAGTTCCTAAGTGTGTTAAGCAAGATGTGTTT 622
Db 182 CAATTGAGAGCACTTACTGAGAGACGCGCCGCCAAGCGTTTCMAAGAGATGTGTGC 241
OY 623 TAGTTATGGGCTTGGCTTCTCGATGTCGAGAGATGCTATGGCAATTTGATCACTG 682
Db 242 TGGCAATGGGCTGCTTATATGTGACATATCAAGGATGCAATGGCGCTTACGCTTCA 301
OY 683 ATTTTATTACTGTTATAGATTTGTTGAGAGAGCTTTGAAGCTTTTACAGAGAGAGAG 742
Db 302 ATGTAATCCGCTGTGTGAGAGTGTGTTGAAGGGCTTCMAAGCTTTTCAAGAGAGATGGG 361
OY 743 CAAGTACCTTGCACCGGATTTTACGTGCAAAATTGATGAGCTTTGGAAGAGATCACTC 802
```

Db	362	CAATCAACCTTGACCTGCTGCTCTTTCACAAATTGATGAAGAACTCTGGAGAGATCACAC	421
Qy	803	CGCGTTATGCTTGGAGCTACTTGGCTTACCGCTTGATGA	842
Db	422	CTCGTTGCTTTTGGAGCTTCTTGCCCTTNCCTTGATGA	461

Search completed: December 10, 2005, 19:21:37
Job time : 1376.27 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 10, 2005, 16:58:05 / Search time 9151.75 Seconds
(without alignments)
12300.344 Million cell updates/sec

Title: US-10-600-070B-1
Perfect score: 2406
Sequence: 1 atggaagctctgagtcacgt.....gtctctgtctctgcataca 2406

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_hic:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	902.2	37.5	919	4	CNS09YXH
2	633	26.3	741	4	CNS09YXH
3	633	26.3	741	4	CNS09YXH
4	482	20.0	561	1	AI998415
5	458	19.0	534	5	BO834167
6	447	18.6	2307	10	CL965374
7	436	18.1	813	7	CO071968
8	429.4	17.8	897	6	CD573714
9	360.8	15.0	785	9	BZ437564
10	359.2	14.3	624	8	CO075595
11	344.2	13.9	769	7	CO075595
12	335.4	13.9	741	8	CX543653
13	328.8	13.7	631	5	BU046755
14	326.6	13.6	703	7	CN918842
15	321	13.3	415	3	BP785511
16	316	12.1	410	3	BP782886
17	306.8	12.8	611	8	CX171798
18	304.4	12.7	341	9	AY199896
19	300.6	12.5	728	8	CX674370
20	299.2	12.4	434	3	BP620404
21	293.4	12.2	832	7	CO079829
22	288.2	12.0	697	8	CX173275

C 23	287.6	12.0	821	7	CK090561	CK090561
C 24	287.6	12.0	922	7	CV264112	CV264112
C 25	286.6	11.9	819	7	CV241483	CV241483
C 26	283	11.8	898	8	DR830109	DR830109
C 27	280.2	11.6	722	7	CO117046	CO117046
C 28	278.8	11.6	660	2	B1268376	B1268376
C 29	272.2	11.3	839	7	CK936162	CK936162
C 30	270.4	11.2	866	6	CD573715	CD573715
C 31	270.2	11.2	723	7	CO079828	CO079828
C 32	265	11.0	746	8	CX674369	CX674369
C 33	261.8	10.9	768	8	CX543654	CX543654
C 34	260.2	10.8	593	1	AM696905	AM696905
C 35	257.4	10.7	574	7	CN904734	CN904734
C 36	254	10.6	849	7	CO117047	CO117047
C 37	249.6	10.4	871	10	DU050213	DU050213
C 38	235	9.8	820	8	DR972351	DR972351
C 39	231.8	9.6	751	11	CR486014	CR486014
C 40	226.6	9.4	521	7	CN914655	CN914655
C 41	226	9.4	683	7	CO076192	CO076192
C 42	224	9.3	537	2	BE490117	BE490117
C 43	220	9.1	501	5	BU049629	BU049629
C 44	218.6	9.1	508	3	BU048486	BU048486
C 45	214.8	8.9	451	1	AJ802091	AJ802091

ALIGNMENTS

RESULT 1
CNS09YXH
LOCUS
DEFINITION
Arabidopsis thaliana Full-length cDNA Complete sequence from clone GSUTS89ZC08 of Adult vegetative tissue of strain col-0 of Arabidopsis thaliana (chale cress).

ACCESSION
BX841670
VERSION
KEYWORDS
SOURCE
ORGANISM
Arabidopsis thaliana
Arabidopsis thaliana (chale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eustroide II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE
AUTHORS
Castelli, V., Aury, J.M., Jallion, O., Wincker, P., Clepet, C., Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V., Temple, G., Caboche, M., Weissenbach, J. and Salanoubat, M.
Whole Genome Sequence Comparisons and 'Full-Length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation

TITLE
Arabidopsis thaliana Full-length cDNA Complete sequence from clone GSUTS89ZC08 of Adult vegetative tissue of strain col-0 of Arabidopsis thaliana (chale cress).

JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
Web : www.genoscope.cns.fr)

COMMENT
The technologies are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Piprap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/full length
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.

FEATURES

source
1. 919
/organism="Arabidopsis thaliana"
/mol_type="mRNA"

gene
/db_xref="taxon:3702"
/clone="GSLTILS92C08"
/tissue_type="adult vegetative tissue"
/ecotype="Col-0"
/plasmid="pCMVSPORT_6"
1..919
/gene="At5g42480"

Query Match 37.5%; Score 902.2; DB 4; Length 919;
Best Local Similarity 99.7%; Pred. No. 4,7e-244;
Matches 904; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

ORIGIN
Query Match 37.5%; Score 902.2; DB 4; Length 919;
Best Local Similarity 99.7%; Pred. No. 4,7e-244;
Matches 904; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
1 ATGAAGCTTGTAGTCAAGTGGGATTTGCTCTCTCCCAATTCATTAATGCCGATTACCA 60
13 ATGAAGCTGTAGTCAAGTGGGATTTGCTCTCCCAATTCATTAATGCCGATTACCA 72
61 CCGGAGAGCAAAAGCTCGAGTGGCCCAACACTCTTACAACTATCTGCTCCGACGC 120
73 CCGGAGAGCAAAAGCTCGAGTGGCCCAACACTCTTACAACTATCTGCTCCGACGC 132
121 AAATGGGCGAGCGCTTCTCTCGACCTTCAATTCACCTCGATTCCCTCCCTCTCC 180
133 AAATGGGCGAGCGCTTCTCTCGACCTTCAATTCACCTCGATTCCCTCTCTCTCC 192
181 TTGCGCACCGCCACCAACCGCCACTCTCTCTCTGCGACCATCTAATGATGATGCC 240
193 TTGCGCACCGCCACCAACCGCCACTCTCTCTCTGCGACCATCTAATGATGATGCC 252
241 GAAGCCGACGCTCCCAATCCCAATTTCTAATTCAGATTTAGAGAGCTTCAAAACATTT 300
253 GAAGCCGACGCTCCCAATCCCAATTTCTAATTCAGATTTAGAGAGCTTCAAAACATTT 312
301 TTAACCGATGATGAGAGAGATTTGAGAGTGGGATTTGAGAGAGCTTGGATTCGAT 360
313 TTAACCGATGATGAGAGAGATTTGAGAGTGGGATTTGAGAGAGCTTGGATTCGAT 372
361 TTGACGCGAGCGCTTATATACGCGGAGACAGATTTCTTCAAGCTTGGCGAACTCTG 420
373 TTGACGCGAGCGCTTATATACGCGGAGACAGATTTCTTCAAGCTTGGCGAACTCTG 432
421 TCTATCTCTGGTCTAAGAGAGATTAAGAGATTTCTTGAATGATGAGAGAGCTACA 480
433 TCTATCTCTGGTCTAAGAGAGATTAAGAGATTTCTTGAATGATGAGAGAGCTACA 492
481 GTATCATCATGATGCTCTTGGGATTAAGGTTCTGGGGCTCTCTGTGTATGCAAGAGAT 540
493 GTATCATCATGATGCTCTTGGGATTAAGGTTCTGGGGCTCTCTGTGTATGCAAGAGAT 552
541 GGTGAGACTGAGATGATTTCTGGGATTAAGGTTCTGGGGCTCTCTGTGTATGCAAGAGAT 600
553 GGTGAGACTGAGATGATTTCTGGGATTAAGGTTCTGGGGCTCTCTGTGTATGCAAGAGAT 612
601 TCGTTTAAGCAAGATGATGTTTAAATGATGATGATGATGATGATGATGATGATGATGAT 660
613 TCGTTTAAGCAAGATGATGTTTAAATGATGATGATGATGATGATGATGATGATGATGAT 672
661 GCTATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
673 GCTATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 732
721 AAGCTTTTAAGCAAGAT 780
733 AAGCTTTTAAGCAAGAT 792
781 GAGATTTTAAGCAAGAT 840
793 GAGATTTTAAGCAAGAT 852
841 GATGATTAAGCAAGAT 900
853 GATGATTAAGCAAGAT 912

Qy 901 GTTGAGAG 907
Db 913 GTTGAGAG 919

RESULT 2

CNS09YNM 741 bp mRNA linear HTC 06-FEB-2004
LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone
DEFINITION GSTIL592A05 of Siliques of strain col-0 of Arabidopsis thaliana
(thale cress).

ACCESSION BX833489.1 GI:42455179
VERSION BX833489.1
KEYWORDS HTC; GSTL cDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana

REFERENCE Castelli, V., Aury, J.M., Jallion, O., Wincker, P., Clepet, C., Menard, M., Cruaud, C., Quetier, F., Scarpelli, C., Schachter, V., Temple, G., Caboche, M., Weissenbach, J. and Salanoubat, M. Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences: A Combined Approach to Evaluate and Improve Arabidopsis Genome Annotation

TITLE Arabidopsis thaliana
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 741)
REFERENCE Genoscope.
TITLES Direct Submission
AUTHORS Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)

COMMENT The sequences are based on single pass reads. Life Technologies (a division of Invitrogen) members carried out full-length libraries construction : Temple G. Genoscope members carried out sequencing and annotation : Castelli V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C., Schachter V., Weissenbach J., Salanoubat M.

URV INRA : Clepet C., Caboche M. Annotation is based on the June 2003 version of the Arabidopsis genome released by MIPS (Munich Information center for Protein Sequences). 5 prime and 3 prime are assembled with Phrap. http://www.genoscope.cns.fr/externe/sequences/Banque_projet_EF/Full length

http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.location/Qualifiers

FEATURES 1..741
Source /organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="GSLTILS92A05"
/tissue_type="Siliques"
/ecotype="Col-0"
/plasmid="pCMVSPORT_6"
1..741
/gene="At5g42480"

ORIGIN
Query Match 26.3%; Score 633; DB 4; Length 741;
Best Local Similarity 100.0%; Pred. No. 1.1e-167;
Matches 633; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1774 GCTGGGGCTCAGAGATGAGTGGTGAAGAACTAGTGAATGTCGTTGATGATG 1833
1 GCTGGGGCTCAGAGATGAGTGGTGAAGAACTAGTGAATGTCGTTGATGATG 60

Qy 1834 TTAAGAGAGCAAGTGTGAAGATCTAGCTGTGTGTGCAATTTGATTTCACTG 1893
Db 61 TTAAGAGAGCAAGTGTGAAGATCTAGCTGTGTGTGCAATTTGATTTCACTG 120

Qy 1894 TTGACCGAAGATTTTCTTAAAGAGCTCATCTTTCAACGCAAGATATGTTCT 1953

Db 121 TTCAGCCAGAGATATTTCTTAAAGACGCTATCTTTTCACGCAAGATATGTTCT 180
QY 1954 TCTATGAATCTGATGTCGCTACCATAGGTCAGTCAAGCTGACGATTCAGAGCACTT 2013
Db 181 TGTATGAATCTGATGTCGCTACCATAGGTCAGTCAAGCTGACGATTCAGAGCACTT 240
QY 2014 CCCAGATGATGCTGATGCTGACGAGAAATATGATTCAGAGTGGCAAGATTAAGTCT 2073
Db 241 CCCAGATGATGCTGATGCTGACGAGAAATATGATTCAGAGTGGCAAGATTAAGTCT 300
QY 2074 CTGGCTTTGGGCGTATCACCGCATAGAAATGTTACGAGAGTTTGGATGGCGCAATG 2133
Db 301 CTGGCTTTGGGCGTATCACCGCATAGAAATGTTACGAGAGTTTGGATGGCGCAATG 360
QY 2134 CTGAAGATTTGGACTGACAGAGACGCTGAAGCTCCGACCTGGTGGTTTATGATTAT 2193
Db 361 CTGAAGATTTGGACTGACAGAGACGCTGAAGCTCCGACCTGGTGGTTTATGATTAT 420
QY 2194 ACACTGTTGAAACTATCTGTTGACAGTGTGACAGTCTCAGCAGATGAAACCGTGTCTG 2253
Db 421 ACACTGTTGAAACTATCTGTTGACAGTGTGACAGTCTCAGCAGATGAAACCGTGTCTG 480
QY 2254 GTGGAAGCAACTCTGAGAGAGTCTGTTGCTATCTGATTTGGTTCATCCAGAAACAT 2313
Db 481 GTGGAAGCAACTCTGAGAGAGTCTGTTGCTATCTGATTTGGTTCATCCAGAAACAT 540
QY 2314 GCTACTGATGTGAGAACTTACACCAAGATGAGAGTTTCTGCTCCAGAGTCAAGGTTG 2373
Db 541 GCTACTGATGTGAGAACTTACACCAAGATGAGAGTTTCTGCTCCAGAGTCAAGGTTG 600
QY 2374 AAATCACTGAAGGCTCTGTTCTTGATCATTA 2406
Db 601 AAATCACTGAAGGCTCTGTTCTTGATCATTA 633

RESULT 3
CNS09255
LOCUS
DEFINITION
CNS09255 741 bp mRNA linear HTC 06-FEB-2004
Arabidopsis thaliana Full-length cDNA Complete sequence from clone
GSLT3IL3Z09 of Silique of strain col-0 of Arabidopsis thaliana
(thale cress).

ACCESSION
BX833051
VERSION
BX833051.1
KEYWORDS
HTC; GSLT_cDNA.
SOURCE
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 741)
Castelli,V., Aury,J.M., Jallion,O., Wincker,P., Clepet,C.,
Menard,M., Cruaud,C., Quetier,P., Scarpelli,C., Schachter,V.,
Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
Whole Genome Sequence Comparisons and 'Full-Length cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
Unpublished
2 (bases 1 to 741)
Genoscope.
Direct Submission
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jallion O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences) . 5 prime and 3 prime are assembled with Phrap.

JOURNAL
REFERENCE
AUTHORS
JOURNAL
COMMENT

http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
length
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsids.
FEATURES
source
1..741
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="GSLT3IL3Z09"
/tissue_type="Silique"
/ecotype="Col-0"
/plasmid="pCMVSPORT_6"
1..741
/gene="At5g42480"
ORIGIN
Query Match 26.3%; Score 633; DB 4; Length 741;
Best Local Similarity 100.0%; Pred.No. 1.1e-167;
Matches 633; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1774 GGTGGGGTCTCAGAGAGTACGCTGATGAACTACGTTGAAATGTCGTTGCTATG 1833
Db 1 GCTGGGGTCTCAGAGAGTACGCTGATGAACTACGTTGAAATGTCGTTGCTATG 60
QY 1834 TTAAAGAGGCAAGTGTGAAGATCTAGCTGCTGTTGGCAATTGACTGATTTCACTG 1893
Db 61 TTAAAGAGGCAAGTGTGAAGATCTAGCTGCTGTTGGCAATTGACTGATTTCACTG 120
QY 1894 TTGAGCCAGAAATATTTCTTAAAGCAGCTCATCTTTCAACGCAAGATATGTTTCT 1953
Db 121 TTCAGCCAGAGATATTTCTTAAAGCAGCTCATCTTTCAACGCAAGATATGTTTCT 180
QY 1954 TCTATGAATCTGATGTCGCTACCATAGGTCAGTCAAGCTGACGATTCAGAGCACTT 2013
Db 181 TGTATGAATCTGATGTCGCTACCATAGGTCAGTCAAGCTGACGATTCAGAGCACTT 240
QY 2014 CCCAGATGATGCTGATGCTGACGAGAAATATGATTCAGAGTGGCAAGATTAAGTCT 2073
Db 241 CCCAGATGATGCTGATGCTGACGAGAAATATGATTCAGAGTGGCAAGATTAAGTCT 300
QY 2074 CTGGCTTTGGGCGTATCACCGCATAGAAATGTTACGAGAGTTTGGATGGCGCAATG 2133
Db 301 CTGGCTTTGGGCGTATCACCGCATAGAAATGTTACGAGAGTTTGGATGGCGCAATG 360
QY 2134 CTGAAGATTTGGACTGACAGAGACGCTGAAGCTCCGACCTGGTGGTTTATGATTAT 2193
Db 361 CTGAAGATTTGGACTGACAGAGACGCTGAAGCTCCGACCTGGTGGTTTATGATTAT 420
QY 2194 ACACTGTTGAAACTATCTGTTGACAGTGTGACAGTCTCAGCAGATGAAACCGTGTCTG 2253
Db 421 ACACTGTTGAAACTATCTGTTGACAGTGTGACAGTCTCAGCAGATGAAACCGTGTCTG 480
QY 2254 GTGGAAGCAACTCTGAGAGAGTCTGTTGCTATCTGATTTGGTTCATCCAGAAACAT 2313
Db 481 GTGGAAGCAACTCTGAGAGAGTCTGTTGCTATCTGATTTGGTTCATCCAGAAACAT 540
QY 2314 GCTACTGATGTGAGAACTTACACCAAGATGAGAGTTTCTGCTCCAGAGTCAAGGTTG 2373
Db 541 GCTACTGATGTGAGAACTTACACCAAGATGAGAGTTTCTGCTCCAGAGTCAAGGTTG 600
QY 2374 AAATCACTGAAGGCTCTGTTCTTGATCATTA 2406
Db 601 AAATCACTGAAGGCTCTGTTCTTGATCATTA 633

RESULT 4
A1998415/c
LOCUS
DEFINITION
A1998415 561 bp mRNA linear EST 08-SEP-1999
thaliana cDNA clone 701545606, mRNA sequence.
ACCESSION
A1998415
VERSION
A1998415.1
KEYWORDS
EST.

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1	Chen, J., Momiyama, M., Chan, E., Mooney, M., Carroon, B., Gilliland, D., Wang, X., Hillman, J., Giegler, K., Kim, C., Doyle, M., Brzozka, P., Gorgone, G., Burns, D., Griffin, J., Mouanoutoua, M., Nguyen, D., Tan, R., Rose, M., Warren, B., Ton, B., Kasbury, K., Borillo, C., Carpio, T., Policky, J., Suzuki, G., Argentine, C., Shah, S., Nobriza, A., Murry, L., Turner, C., Krikorian, S., Elder, L. and Hanson, D.	Arabidopsis thaliana Gene Expression Microarray	Unpublished (1999)	
2	Genome Systems, Inc., a wholly owned subsidiary of Incyte Pharmaceuticals, Inc.	4633 World Parkway Circle, St. Louis, MO 63134, USA	Tel: 877-577-2733	
3	Fax: 314-427-3324	Email: serv@genomesystems.com.	Location/Qualifiers	
4	1. .561	/organism="Arabidopsis thaliana"	/mol_type="mRNA"	
5	/ecotype="Col-0"	/db_xref="taxon:3702"	/clone="701545606"	
6	/issue_type="rosette"	/dev_stage="4 - 7 weeks"	/clone_lib="A. thaliana, Columbia Col-0, rosette-2"	
7	/note="vector: pSPOR, Site 1: NotI, Site 2: SalI, cDNA library was derived from untreated rosette tissue from Arabidopsis thaliana, Columbia Col-0, at 4 - 7 weeks. Plants were grown in 1:1:1 peat moss/vermiculite/perlite soil at 22 deg. C +/- 3 deg. C under constant light, and watered with fertilizer. cDNA synthesis was initiated using a NotI-oligo(dT) primer. Double-stranded cDNA was blunt-ended, ligated to SalI adaptors, digested with NotI, size-selected, and cloned into the NotI and SalI sites of the pSPORT vector."			
8	ORIGIN			
9	Query Match	20.0%; Score 482; DB 1; Length 561;		
10	Best Local Similarity	99.4%; Pred. No. 7.5e-125;		
11	Matches 482; Conservative	0; Mismatches 3; Indels 0; Gaps 0;		
12	QY	1922 GCTCATCTTTTCAACGACGAGATATGATGTTCTTCTATGAGATGATGTGCTACCATAG	1981	
13	DB	561 GCTCATCTTTTCAACGACGAGATATGATGTTCTTCTATGAGATGATGTGCTACCATAG	502	
14	QY	1982 GGTGAGTCAGAGCTGACGATTCAGAAAGCACTCCGACGAGATGATGCTAGAGCTGACAGAG	2041	
15	DB	501 GGTGAGTCAGAGCTGACGATTCAGAAAGCACTCCGACGAGATGATGCTAGAGCTGACAGAG	442	
16	QY	2042 ATATAGATCCAAATGCGACAAAGATTAAGTCTCTGGCTTTTGGGCTGATTCACCGCATAG	2101	
17	DB	441 ATATAGATCCAAATGCGACAAAGATTAAGTCTCTGGCTTTTGGGCTGATTCACCGCATAG	382	
18	QY	2102 AAATGTTACCAAGAGTTTGGATGGGCAATGCTGAAAGATTGGAATGACACAGACACTG	2161	
19	DB	381 AAATGTTACCAAGAGTTTGGATGGGCAATGCTGAAAGATTGGAATGACACAGACACTG	322	
20	QY	2162 AAATGCGGAGCTTGGGTTTATATGATTAATTAACCTGTTGAAAATATCTGTTGACAGTG	2221	
21	DB	321 AAATGCGGAGCTTGGGTTTATATGATTAATTAACCTGTTGAAAATATCTGTTGACAGTG	262	
22	QY	2222 TGAACGCTCAGCAGATGGAACCCGCTGCTCGGTGGAAGCAACTCTGGAAGAGTCTGCTT	2281	
23	DB	261 TGAACGCTCAGCAGATGGAACCCGCTGCTCGGTGGAAGCAACTCTGGAAGAGTCTGCTT	202	
24	QY	2282 GTCTATCTGATTTGGTTTCATCCAGAAAAACATGCTACTGATGTGAGAACTTACACACAA	2341	

Db	201	GTGATCTGATTTTGTTTCATCCAGAAAACAATGCTACTGATGTCAGAACTTACACAA	142					
Qy	2342	GATACGAAGTTTCTGTCCTCAAGTCAGGTCGAAAATCACTGAAGGCTCTGTTCTTGCA	2401					
Db	141	GATACGAAGTTTTCGTTCCAAAGTCAGGTCGAAAATCACTGAAGGCTCTGTTCTTGCA	82					
Qy	2402	CATPA 2406						
Db	81	CATPA 77						
RESULT 5								
LOCUS	B0834167	534 bp	mRNA linear EST 14-MAR-2003					
DEFINITION	ABEST0128 Arabidopsis lyrata Inflorescence pCMV-PCR Library							
ACCESSION	B0834167	Arabidopsis lyrata cDNA clone PIWB1-D03 3', mRNA sequence.						
VERSION	B0834167.1	GI:28951482						
KEYWORDS	EST							
SOURCE	Arabidopsis lyrata							
ORGANISM	Arabidopsis lyrata							
REFERENCE	Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.							
AUTHORS	Barlier,M., Bustamante,C.D., Yu,J. and Putuganan,M.D.							
TITLE	Selection on rapidly evolving proteins in the Arabidopsis genome							
JOURNAL	Genetics 163 (2), 723-733 (2003)							
PUBMED	12618409							
COMMENT	Contact: Barrier M Department of Genetics North Carolina State University 3513 Gardner Hall, Box 7614, Raleigh, NC 27695, USA Tel.: 919 515 1761 Fax: 919 515 1695 Email: mbarlier@unity.ncsu.edu Plate: 1 row: H column: 5 Seq primer: T3 Location/Qualifiers 1..534 /organism="Arabidopsis lyrata" /mol_type="mRNA" /cullivar="Kathumaki" /db_xref="taxon:59689" /clone="PIWB1-D03" /tissue_type="Inflorescence" /clone_id="Arabidopsis lyrata Inflorescence pCMV-PCR library" /note="Vector: pCMV-PCR (Stratagene); Created using PCR library construction kit (Stratagene)"							
FEATURES								
Source								
Query Match 19.0%; Score 458; DB 5; Length 534; Best Local Similarity 94.1%; Pred. No. 4,8e-118; Matches 476; Conservative 0; Mismatches 30; Indels 0; Gaps 0;								
ORIGIN								
Qy	1901	AGAAATATTTTCTTAAAGCAGCTCATCTTTTCAAGCAAGATATGTTCTTCTATGG	1960					
Db	2	AGATTTTGCTCTCAAAAGCAGCTCATCTTTTCAAGCAAGATATGCTTCTTCTATGG	61					
Qy	1961	AATCTGATGTCGTACCATGAGGTCAGTCAGCTGACATTCAGAAAGCACTTCCAGAA	2020					
Db	62	AATCTGACATTTGCTTCCATGGGTCAGTCAGCTGATTCAGAAAGCACTTCCAGAA	121					
Qy	2021	TGATGCTTACGACTGCGAGAAATATATTCACAGTCGCGAAGATTAAGTCTCTGGCTT	2080					
Db	122	TGATGCTTACGACTGCGAGAGATATATTCACAAAGTCGCGAAGATCAAGTCTCAGGCTT	181					
Qy	2081	TTGGGCTGATCAACCGCATAGAAATGTTTACAGAGGTTTGGATGGCGAATGCTGAAGA	2140					
Db	182	TTGGGCTGATCACTGATGAAGATGTTTACAGAGGTTTGGATGGCGAATGCTGAAGA	241					

QY 2141 TTGGAATGACAGAGAGCTGAAACTGCCAGCTTGGTTGTTATGATTAATACCTGT 2200
Db 242 TTTGACCTGACAGAGCTGTGGAAACCTGCCAGCTTGGGCTGGTTATGATTTATACCTGT 301
QY 2201 TGAACCTATCTGTGACAGTGTGACAGTGTGACAGATGGAACCGTGCTGTGGTGAAG 2260
Db 302 TGAACCTATCTGTGACAGTGTGACAGTGTGACAGATGGAACCGTGCTGTGGTGAAG 361
QY 2261 CAACTCTGAGAGAGCTGTGCTGTCTATCTGATTTGGTTGATCCAGAAAACAATGCTACTG 2320
Db 362 CAACTTTGAGAGAGCTGTGCTGTCTATCTGATTTGGTTGATCCAGAAAACAATGCTACTG 421
QY 2321 ATGTGAGAACTTACACAAAGATGCGAAGTTTCTGTGTCAGTCAAGGTGGAATATCA 2380
Db 422 ATGTGAGAACTTACACAAAGATGCGAAGTGTGTCAGTCAAGGTGGAATATCA 481
QY 2381 CTGAAGGCTGTGCTGTGATCATTA 2406
Db 482 CTGAAGGCTGTGCTGTGATCATTA 507

RESULT 6
CL965374 2307 bp DNA linear GSS 21-SEP-2004
LOCUS CL965374
DEFINITION OsJFCC012338 Oryza sativa Expressed Library Oryza sativa (indica
cultivar-group) genomic, genomic survey sequence.
ACCESSION CL965374
VERSION CL965374.1 GI:52385433
KEYWORDS GSS.
SOURCE Oryza sativa (indica cultivar-group)
ORGANISM Oryza sativa (indica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
1 (bases 1 to 2307)
Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M.,
Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L.,
Wong, G. K. S., Deng, X. W. and Wang, J.
An analysis of transcriptional regulation of the rice genome and
its comparison to Arabidopsis
Unpublished (2004)
JOURNAL
COMMENT Contact: Chen Chen
Department of Bioinformatic
Beijing Institute of Genomics
Chinese Academy of Sciences, Beijing 101300, China
Tel: 86-10-80481559
Fax: 86-10-80488676
Email: chenchen@genomics.org.cn
Rice genomic sequence.
Class: exon-trapped.
FEATURES
source 1..2307
Location/Qualifiers
/organism="Oryza sativa (indica cultivar-group)"
/mol_type="genomic DNA"
/db_xref="taxon:39946"
/clone_id="Oryza sativa Expressed Library"
/note="Oryza sativa exon trapped genomic sequences"

ORIGIN
Query Match 18.6%; Score 447; Db 10; Length 2307;
Best Local Similarity 53.7%; Pred. No. 1e-114;
Matches 1256; Conservative 0; Mismatches 880; Indels 201; Gaps 8;
QY 108 CTGCTCCGACGAATAGGCGACCGCTTCTCTCGACTTCAATTTACCTCGATTC 167
Db 132 CTGGGCGGAAAGCCCTTTCGCGACTTCCACTCTCTCCCAACGCGGCGCTTCGACCC 191
QY 168 CTCTCTCTCTCTTTCGACCGACCAACGCACTCTCTCTCTCTCTGCAACATC 227
Db 192 GCGGCGCGGCTT 251
QY 228 TAATGATGCTCCCAAGCGCAAGTCCCATCCCATTAATTTCTACAGGAT----- 280

Db 222 CCCCAGCGCCCGAAACGCTCCCTCCGCTCCAAAGTCAATTTCTAACAAAGTAGGCCCCG 311
QY 281 -----TAGAGCTCAACACATTTCTTAACCGATGAAATCAGAGAGC 323
Db 312 TGAATCTGCTGCGTTCTTAGGGGAGAGACACATTTCTTGGCGATGGCATAGAGAGGC 371
QY 324 ATTGAAAGCTAGGGTTTGAACCGCGCAATTGCTTTCAGCGACGACGCTTTAATCAG 383
Db 372 GTTCAGAGCAGGATAGCAGCAACCGCAGATAGGCTACAGCAGAGATGCTCTTGTTGG 431
QY 384 CCGAGACAGATTTCTTCAAGCTGCTTGGAAACTCTGTCTAATCTCGGTCTAGAAAGA 443
Db 432 TCGTCGCAAAATGCTGCGAATGCTCCATGACCTCTCATGAAACCAAACTCCGACATCA 491
QY 444 GTACAAATGAAGTCTTCTTGATGATGAAGAGCTACAGTCACTCATGATGTTCTTGGGA 503
Db 492 GTATGATGTGCGCTTCTTGAGAACCGGAAAGAGCTTCAACATGATATTTGCTTGGGA 551
QY 504 TAAAGTTCTTGGGCGCTCTGTGTATGTCAGAAAGGTGAGACTGAGATAGTCTTTCG 563
Db 552 CAAGGAGGCTGGG-----GAGGCACTTGTGCTGTGT 584
QY 564 GGTGTGAGAGCTCTGCTTAAGAGAGAGTTCCTTAAGCTTTAAGCAAGTGTGTTT 623
Db 585 AACTGAGAACAGTTGCTTCTGATGCGCACCAAGGCTTCAAGCAGAGAGTGTGCT 644
QY 624 AGTTATGCGCTTGTGCTTCTGATGTCGAGAGGATCTATGAGCATTTGATTCACCTGA 683
Db 645 AGCGATGCTTGTGCTTATGTGATCTATCAAGGATCTATGCGACAGACCTTCCAGA 704
QY 684 TTTTATTAATGCTTATGATGATTTGTTGAGAAAGCTTTGAAGCTTTTACAGAGAAAGAG 743
Db 705 TGTATATGCTGCTGCGAGAGGCTCTGAGAGGCTCTCAAGCTCTTGCAGAAATGAGAGC 764
QY 744 AAGTAGCTTGACACCGGATTTACGTGACAAATTTGATGAGCTTTGGAAGAGATCATCC 803
Db 765 AAGCAATCTGCGACCTGATGCTTTCACAGATTTGAAACCTCTCGAGAGATTTACACC 824
QY 804 GCGTTATGCTTGGAGCTACTTGTGCTTACCGCTTGTGATGATTAAGCTGGAAGAAAGCT 863
Db 825 TCGCTGTATGATGAGCTTCTCTCCCTTCTTATGACACAGAGATCATTAAGAACGCGCA 884
QY 864 AATAGTTTAAAGCGGTGCGGAATATTTGTGCTGTGAGAGAGTGAAGCATCAGC 923
Db 885 AGAAGGCTTCAAGGTGAGAAACATTTGTGAGGTTGGCAGAGAGGATATTGCTAC 944
QY 924 TCTTGTGGGGGTTTGAACCGCTGAGAAAGTTATGATGAGAGCGCTTTTTCAGATGACAGC 983
Db 945 CGTGTGAGAGGATTTCTCGTGAAGCTTCATGAAGAGCTTTTGTGAGAGATGACATC 1004
QY 984 TGCTGACAGGTTGATCTTTTGTAGCTACCCCAAGCAATTTCCAGAGAGTCATTTGA 1043
Db 1005 AATTGAACAGATGATTTCTTTTCAAAAACCGAATGACATTCCTCTGATAGGTTTGA 1064
QY 1044 AGTTTAAAGTTGACCTTGTCTTGTGCTCAAGCTTTTATTTGTAAGAACCAACCT 1103
Db 1065 AATTTAACAATGACACTTGTGACATGTGCTCAAGCAATTAATGATTAAGAGCCACATTT 1124
QY 1104 TTTTACAGATGCTATTAAGCAATTTCCAGCACTTCAAGAGCTTAAGTATGCTATGGA 1163
Db 1125 CATATATATGCGGATATCTTTTGAACAATCCAAAGTTCA----- 1169
QY 1164 GATTCCTGCAATGTTATGATACAGGAATTAATTTGGAAGATAGACTTGGCTAGAAAG 1223
Db 1170 -----CATAGGTTCTATTATGCTTATGATTAATGAGATGAGACTTGTGATGGAAG 1220
QY 1224 GGAATCTGTGACATGCTTATTAAGCAAAATGATGAATGCGATATGTTGGGCTTGA 1283
Db 1221 GGCAATCTGCTCATTTGCTAGCGAGATGTTAGCAATGACGAATGTGGCTTGAATTTGA 1280
QY 1284 CAGTGAAGATTTCAAAATATGAGATTCAGATATTTGTGAGATTTCTTTTGAAGATTC--- 1340
Db 1281 TAATGATCTTCAACATACAGAGACCCCAAAATTTTATGAGTTATTTGTGACCAACTCTAG 1340

1341 AAATGATGACATGATGATCTCCCTGACTATGCAATTTGTTGAAACCTGTTGGC 1400
1341 CATGAGTAAGATGATCTCTCCAGGGCTGTGCAAGCTTTTGAGACTTGGCTTAT 1400
1401 AGGGTTCTCTTCTAGGTTCAAGACCAAAAGATTAATAAATTTAACTCGGGACTA 1460
1401 CTTTGGGTTTCTTAGAGACAGATATCTGGGGCATGTGACACTTGAAGATTA 1460
1461 CTATGATGATCTTATGTTTGAATTAAGTGAAGAGTGAAGTTCAGGGTTCTCC 1520
1461 CTAGCATATCAAGAGTTTAAAGTCTTGAAGAGATGAGGGTGTGCTCTTCA 1520
1521 TTTAGCTGCTCTGCACTATGCAAGGATTTGAGCCGACAGATGAAAGTACTAT 1580
1521 TTTGGCTCTGCTGCTATTTGCAAACTGTGCTCAGCTACAGCTGACCTTGGTA 1579
1581 GCAGGCACTGCAAGAAATTTTCTTCCCTGCTATACAGATAGAACTGCTGAACCA 1640
1580 -----CTGGAATCAAA 1592
1641 GGATGTCAGAGACAGTTTATGATGATCTGTTGTAAACAATGAGCCGTGATCG 1700
1593 TGCTATTCAGCG----- 1605
1701 TGAGCTGTGTTCTTTATGCAAGCTGTAAAGCCCTTGAAAACTTTGAACATAA 1760
1606 -----TTCAACAAGTTTTCATGATGATTAACAAGTTAGACAGGTGACG 1649
1761 TTTATGCAATTCAGAGCTGGGTTCTCAGAGATGAGCTGTATGATAAACTGTTAAATCTC 1820
1650 CATGAAAAATCTAAAGATGCGCCCTGGGGGATATCTTGAATAATTTTGAACGAAATGCG 1709
1821 CGTTGCTGATATGTTAAAGAGGCAAGTGAAGTCTTACTGCTGCTGTGGAATTTGG 1880
1710 ACTGCTATGATTCGAGAAATGCGCCTTGAAGATTAATCTGCTGCGCACTGTTTGC 1769
1881 ACTGATTTCACTGTTCAACGAGAAATTTTCTTAAAGACGCTATCTTTCAACGCA 1940
1770 ACTGTGCGAGTAATTTGGGGCAAAATTTGCTGTAAAGAGGCCCTTTGCTATTAAG 1829
1941 GGATATGTTCTTCTTATGAAATCTGAT-----GTGCTACCATAGGCTCAGTCAG 1991
1830 GAGTGAAGATGATCTGTGCGAGTTGCTAATATGATCTCACTCTGATGATCTGCACT 1889
1992 AGCTGACGATTCAGAAAGCACTTCCAGAAATGAGATCTGAGACTGCGAGAAATATAGTATC 2051
1890 AGATGAAGATCCAGTACATATTTCTGAAATGATGCGAAGCTGGCAGAAATATTTGTCG 1949
2052 CAACTGCGAGAAATTAAGTCTTGGCTTTGGGCTGTATCAACGCAATGAAATGTTACC 2111
1950 CAACTGCGAGAAATTAAGTCTTGGGCTTGGGACCAAGATTCGCTGATCATTTGCA 2009
2112 AGAGTTTGGATGCGGCAATGCTGAAGATTTGCACTGACAGAGAGTGAATGCGCA 2171
2010 AGAGTTTCTTATGCAACATGCTTAAAGTGTGACTGACGAGAGAGGAGATTTGAACG 2069
2172 GCTTGGTGTGTTTATGATTAATCACTGTTGAACTATCTGTTGACATGAGAGTCTC 2231
2070 TCAATGGTGTGTTCTGGAGATTAACATCCGATGTGACATGATGATCACTATCTC 2129
2232 AGCAGATGAACCCGCTGCTGTGAGAGCAACTCTGAGAGAGTCTGCTGTATCTGA 2291
2130 CCTAGATGTTGAGAGAGAGCTGTGAGAGCTAGATGATGAGAGAGGCAACTTACTGA 2189
2292 TTTGTTTATCCAGAAACATGCTACTGATGCTCAAGCTTACACAGACAGATGAGAGT 2351
2190 TGTTACTAGGCCAGAAACATGATTCATATGACAAATAATCACTTACCCGATATGAT 2249
2352 TTTCTGCTCAGAG---TGAGGTGAGAAATCACTGAAGGCTGTTCTTGTGATCATA 2405
2250 GAGCTTCTCAGAGCTGAGAGGTTGAGAAATTAAGAGAGAGAGTCTCAAGTCTGA 2306

RESULT 7
CO071968 813 bp mRNA linear EST 15-JUN-2004
LOCUS CO071968
DEFINITION GR_Ea30N03.r GR_Ea Gossypium raimondii cDNA clone GR_Ea30N03.3',
mRNA sequence.
ACCESSION CO071968
VERSION CO071968
KEYWORDS GI:48741449
SOURCE EST.
ORGANISM Gossypium raimondii
Gossypium raimondii
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
1 (bases 1 to 813)
REFERENCE
AUTHORS Kam,H., Yu,Y., Kudrna,D., Hatfield,J., Stum,D., Mueller,C.,
Udall,J.A., Rapp,R.A., Wende,J.F., Rao,K., Soderlund,C. and
Wing,R.A.
TITLE Global assembly of Cotton ESTs
JOURNAL Unpublished (2004)
COMMENT Contact: Rod A. Wing
Arizona Genomics Institute
The University of Arizona
Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: rwing@genome.arizona.edu
Plate: 30 row: N column: 03.
FEATURES
source
location/Qualifiers
1..813
/organism="Gossypium raimondii"
/mol_type="mRNA"
/db_xref="taxon:29730"
/clone="GR_Ea30N03"
/tissue_type="whole seedlings"
/dev_stage="first true leaves"
/lab_host="DH10B"
/cnote="Vector: PCMV.SPORT-6.1; Site 1: NotI; Site 2:
EcoRV; Library made by Invitrogen with RNA supplied by
Wendle lab. Directional cloned into NotI-EV. Clonones
plated/picked by AGT. More glycerol clones held in -80."

ORIGIN
Query Match 18.1%; Score 436; DB 7; Length 813;
Best Local Similarity 72.3%; Pred. No. 9,7e-112;
Matches 584; Conservative 0; Mismatches 215; Indels 9; Gaps 1;
Query 107 TCTGCTCGGCGAGAAATGCGGCGGACCGCTTCTCTCCGACTTCAATTTCACTCG--- 163
Db 5 TCTGCTCGGCGAGAAATGCGGCGGACCGCTTCTCTGACTTCAATTTCTCGGCGCC 64
164 -----ATTGCTCTGCTCTCTCTGCGGCGGACCGGACCAACCGGCACTCTGCTCTC 217
Query 218 TGCCACCATTAATGATCGTCCGAAAGCCAGCGTCCCATCCCATTTGATTTTCAACAG 277
Db 65 CCGATTAATTCGTTCTCTCTTCTCTCTCAACCGGCACTTTTCTCCCTTAACTCC 124
125 CCGTCTCGGCGGCTCTCTCTCTCCGAGGCGGACGTTTCAATTTCCCTTGAAG 184
278 TATTAGAGCTCAACACATTTCTTAAACGATGATCAGAGAGATTTGAAAGCTAGG 337
185 TTTTGAAGGCGGAGACTCATTTCTTATGATGATGAAATCAGAAAGGCTATGAAGAGG 244
Query 338 TTTTGAAGGCGGCGGCAATTTGTTGAGGAGAGAGCTTAAATCAAGCGGAGAGATTC 397
Db 245 TTTTGAAGGCGGCTCAATATGATGATGATGATGATGATGATGATGATGATGATGAT 304
398 TTTCAAGCTGCTTGGAAACTCTGCTAATCTGCTGCTTGAAGAGAGATCAATGAAGTCT 457
Db 305 TTTCAAGCTGCTGAGAAACCTTATCTAACCTGCTCTTAACGAAATTAACACCAAGTCT 364
Query 458 TTTTGAATGATGAAGAGCTACAGTATCACTGATGTTCTTGGGATGAAGTCTCTGGG 517

```

Db      365 TTGTCAGCATGAGCGGTGACATATCATCTCACTCAAGTCCCTGGAGACAAGGTTCTTGAG 424
Oy      518 CTCTCTGATATGCAAGAGGTTGGTGAAGTGAAGTATGTTCTTGCGGTTGGTGAAGCTC 577
Db      425 CATTCGTGCTGTGCAAGAGCTGGGAGACTGAGTGTGCTTCAATTTGGGAGAGTT 484
Oy      578 TGCTTAAGAGAGGTTGCTTAAGTCTTTAAGCAAGATGTGTTTAAATTATGAGGCTTG 637
Db      485 TGCTGAGAGAGGCTGCTCCCAAGCATTTTAAACAGAGTGTGCTTGCAATGAGCCCTG 544
Oy      638 CGTTCTGATGCTCTGAGAGGATGCTATGAGCTTGAATCCATTTTATATGCTGT 697
Db      545 CTATGTCGACTTTCGAGGAGTGTATGCTTTGATCTTCACATTTTCATTTGCTGT 604
Oy      698 ATGAGTTGTTGAGAGACCTTTGAAGCTTTTACAGAGAGAGCAAGTACCTTGCAC 757
Db      605 GTGAGTCTCTGAGAGGCTTTGAAGCTTTGACAGAGAGAGTCCAGCACTTGTCTC 664
Oy      758 CGGATTTACGTGCACAAATGATGAGACTTTGGAAGAGATCATCCGCTTATGCTTGG 817
Db      665 CGGATTTACGATCACAGATGATGAGACTTTGAAGAGATTTACCCACGTTGTGCTGG 724
Oy      818 AGCTACTGCTTACCCCTTGTGATGATGATGATGATGATGATGATGATGATGATGATG 877
Db      725 AACTTCTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTG 784
Oy      878 GGTGCGGATATTTTGTGCTGCTTGG 905
Db      785 GTGTAGCGACATATTATGAGGCTGTGG 812

```

```

RESULT 8
CD573714      897 bp      mRNA      linear      EST 12-JUN-2003
LOCUS        UCRPT01_01_F12_T3 Poncirus trifoliata CTV-challenged cDNA library -
DEFINITION   UCR Poncirus trifoliata cDNA clone UCRPT01_01_F12, mRNA sequence.
ACCESSION   CD573714
VERSION     CD573714.1 GI:3169616
KEYWORDS     EST.
SOURCE       Poncirus trifoliata
ORGANISM     Poncirus trifoliata
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Sapindales; Rutaceae; Poncirus.
1 (bases 1 to 897)
REFERENCE    1. Roose, M.L., Ye, X., Federico, C.F., Close, T.J., Fenton, R.D.,
AUTHORS      Wamamaker, S., Choi, Y. and Kingan, T.
TITLE        Development of EST Resources and New Genetic Markers for California
JOURNAL      Citrus - Poncirus trifoliata CTV-challenged phloem - UCR
COMMENT      Unpublished (2003)
CONTACT      Contact: Mikeal Roose
DEPARTMENT   Department of Botany & Plant Sciences, University of California
LOCATION       Riverside, CA, 92521-0124, USA
TELEPHONE    Tel: 909/7874137
FAX          Fax: 909/7874437
EMAIL        Email: mikeal.roose@ucr.edu
SEQUENCE     Seq primer: T3.
FEATURES
SOURCE
1. 897
Location/Qualifiers
/organism="Poncirus trifoliata"
/mol_type="mRNA"
/cultivar="Pomeroy OP"
/db_xref="taxon:37690"
/clone="UCRPT01_01_F12"
/tissue_type="Phloem"
/dev_stage="10 - 30 cm shoots"
/lab_host="E. coli TJC121"
/clone_lib="Poncirus trifoliata CTV-challenged cDNA
library - UCR"
/note="Vector: Lambda Uni-ZAP XR, excised phagemid;
Site 1: EcoRI, Site 2: XhoI. Plants were grown in the
greenhouse at University of California, Riverside. The

```

ORIGIN

scion was a open-pollinated (very probably selfed) seedling of Poncirus trifoliata cv Pomeroy that was selected as homozygous for the CTV resistance gene. The rootstock was sweet orange infected with citrus tristeza virus (CTV) isolate T514 over 1 year before sampling. CTV infects sweet orange, but not genotypes carrying the CTV resistance gene. Shoots 10-30 cm long were harvested in October 2000, and the green phloem (bark) was removed and frozen quickly in dry ice. Total RNA was extracted using Trizol reagent (Gibco). Poly(A) RNA was purified, a cDNA library was made, and 0.5 million primary lambda cDNA clones were in vivo excised to give a population of phagescript SK(-) phagemids. All steps to this point were performed in the ML Roose lab at the University of California, Riverside by X. Ye. Phagemids were plated, plasmid DNA purified, cDNA clones archived, and DNA sequences determined bi-directionally using an ABI730 at the University of California Riverside Genomics Institute, Core Instrumentation Facility, (Choi, Kingan). Chromatogram files were transmitted to UC Riverside (by Choi), then processed at UC Riverside (by Wamamaker) using the HarvesT pipeline (<http://harvest.ucr.edu>) to remove vector and cloning oligo sequences and various contaminants, and to trim to a high quality region. Sequences that retained a phred 17 region of at least 100 bases were deposited to Genbank."

```

Query Match      17.8%; Score 429.4; DB 6; Length 897;
Best Local Similarity 70.4%; Pred. No. 7.5e-110;
Matches 574; Conservative 0; Mismatches 241; Indels 0; Gaps 0;

```

```

Oy      742 GCAAGTACCTTGCACCGATTTTACGTGCACAAATGATGAGCTTTGGAAGATCACT 801
Db      80  GCCAGTACTCTGCTCCAGATTTGACGACAGATGACGAGTGGAGAGATCAAC 139
Oy      802 CCGGCTTATGCTTGAAGCTTACCTGCTTACCGCTTGTGATGATTCGCTCGGAAAGA 861
Db      140 CCACGCTGTGTTTGGAACCTTTTACCTTGTGTGATGATGATGATGATGATGATGATG 199
Oy      862 CTAAATGTTTAAACGGTGGGGAATATTTTGTGCTGTTTGAAGAGGTGGAGCATCA 921
Db      200 GAAAGAGACTTCATGATGATCTCAACATGATGATGATGATGATGATGATGATGATG 259
Oy      922 GCTCTTGTGGGGGTTTGAACCGTGAAGTTTATGATGAGGGCTTTTACGAATGACA 981
Db      260 GCATTTGTGGGGGATTTACAGCTGAATCTTTCATGATGAGGGCTTTTACGATGACA 319
Oy      982 GCTGCTGACAGGTTGATCTTTTGTGATGATGATGATGATGATGATGATGATGATG 1041
Db      320 TCAGCTGAGCAGGTTTAACTATTTTGTGCTGACCAACCAATATGATGCCAGTGAACCTTT 379
Oy      1042 GAACTTACGAAGTTGACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1101
Db      380 GAACTTATGAGAGTGACCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 439
Oy      1102 CTTTACAGATGCTGATTAAGCAATTCAGCACTTGCAGAGCTTATGATGATGATGATG 1161
Db      440 CTGATGAGATGCTGATTAAGATTTCAAGCACTTGCAGAGATTAAGATTAAGCACTTTA 499
Oy      1222 AGGGGACTCTGTGACCTGCTTATGAGCAAGTTGATGATGATGATGATGATGATGATG 1281
Db      560 AGGGGCTGATGTTTACCTGCTGATGATGATGATGATGATGATGATGATGATGATG 619
Oy      1282 GACAGTGAAGTTCACATATAGGAATCCAGCTATGATGATGATGATGATGATGATGATG 1341
Db      620 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 679
Oy      1342 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1401

```

Db 680 AAGGAAGCTAATGACATGATCTTCTGATCTCTGTAAGCTGTAGACATGCGTGCAC 739
Qy 1402 GGGGTTCCTTCTCCAGTTGACAGACACCAAGTATAAAATTTAACTGGGACACAC 1461
Db 740 GAGGTGTTTCTCCAGTTGACAGACACCAAGTATAAAATTTAACTGGGACACAC 799
Qy 1462 TATGATGATCTTATGATCTTCTGATCTTCTGATCTTCTGATCTTCTGATCTTCTGAT 1521
Db 800 TATGATGATCTTATGATCTTCTGATCTTCTGATCTTCTGATCTTCTGATCTTCTGAT 859
Qy 1522 TTAGCTGCTGCTGCAACTATGCAAGATTGAGC 1556
Db 860 TTAGCTGACAGCTGACCATAGATTAGCTTGAGC 894

RESULT 9
B2437564/c B2437564 785 bp DNA linear GSS 13-DEC-2002
LOCUS BONRN72TR.B0.1.6.2 KB tot Brassica oleracea genomic clone BONRN72,
DEFINITION genomic survey sequence.
ACCESSION B2437564 GI:26691135
VERSION B2437564.1 GI:26691135
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 785)
Aylee, M., Haas, B.-J., Kumar, N., Wu, H., Xiao, Y., Van Aken, S.,
Utecherbeck, T.R., Wortman, J.R., White, O.R. and Town, C.D.
Whole genome shotgun sequencing of Brassica oleracea and its
application to gene discovery and annotation in Arabidopsis
Genome Res. 15 (4), 487-495 (2005)

JOURNAL
PUBMED 15805490
Other GSSs: BONRN72TR
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TR
Clas: sheared ends.

FEATURES
Location/Qualifiers
source 1..785
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="T01000DH3"
/db_xref="taxon:3712"
/clone="BONRN72"
/clone_1lb="B0.1.6.2 KB tot"
/note="Vector: PHOS1; Site 1: BactXI; 1.6-2 kb sheared
total DNA inserted into PHOS1 using BactXI linkers"

ORIGIN
Query Match 15.0%; Score 360.8; DB 9; Length 785;
Best Local Similarity 86.4%; Pred.No.2:1e-90;
Matches 425; Conservative 0; Mismatches 57; Indels 10; Gaps 2;

Qy 505 AAGGTTCCTGGGAGCTCTGATGATGCAAGAGGTGAGACTGAGATAGTTCTTCGG 564
Db 633 AAGGTTCCTGGGAGCTCTGATGATGCAAGAGGTGAGACTGAGATAGTTCTTCGG 634
Qy 565 GTTGTGAGGCTCTGCTTAAGAGAGGTTGCTTAAGCTGTTTAAGCAAGATGCGTTTAA 624
Db 633 GTAGAGAGAGCTTCTGCTTAAGAGAGGTTGCTTAAGCTTCAACAGATGCGTTTAA 574
Qy 635 GTTATGAGGCTTCTGCTTCTGATGCTGAGAGATGCTATGAGCTTCCACTTAT 684
Db 573 GTTATGAGGCTTCTGCTTCTGATGCTGAGAGATGCTATGAGCTTCCACTTAT 514

Qy 685 TTATCTAGTGTATGAGTTGTTGTTGAGAGAGCTTTGAGCTTTTACAGAGAGAGCA 744
Db 513 TTTATCTAGTGTATGAGTTGTTGTTGAGAGAGCTTTGAGCTTTTACAGAGAGAGCA 454
Qy 745 AGTACCTTGACCGGATTTAGCTGACCAATGATGAGACTTTGAGAGATCATCTCG 804
Db 453 AGCAGCTTGACCGGATTTAGCTGACCAATGATGAGACTTTGAGAGATCATCTCG 394
Qy 805 C-GTTATGCTGAGGCTTACCTGCTTACCGCTTGAGATGATTAAGCTGCAAAACACT 863
Db 393 CGGTATGATGATGAGCTTCTGCTTACCTTGGGGGAATGAT-----AAACACA 343
Qy 864 AATGTTTAAAGCGGTGCGGAATATTTGTTGCTGTTGAGAGAGTGAAGCATCAGC 923
Db 342 AGATGTTTAAAGCGGTGCTTACGCAATATCTGTGCTGTTGAGAGAGCGGAGCATCAGC 283
Qy 924 TCTTGTGGGGGTTTGAACCGGTGAGAAATTGATGAGAGCGCTTTTACGAATGACAGC 983
Db 282 TGTGTTGGTGTGCTTACAGGTGAGAAATTTATGATGAGAGCGCTTTTACGAATGACAGC 223
Qy 984 TCGTACGACAGT 995
Db 222 ATTCAGCAAGT 211

RESULT 10
CX195205 624 bp mRNA linear EST 28-DEC-2004
LOCUS 13-E019365-021-009-J03-TR ADIS-MP12 021 Brassica napus cDNA clone
DEFINITION MP12p102J0039Q, mRNA sequence.
ACCESSION CX195205
VERSION CX195205.1 GI:56842629
KEYWORDS EST.
SOURCE Brassica napus (rape)
ORGANISM Brassica napus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 624)
Jakoby, M., Lehmann, D. and Weisshaar, B.
direct submission to Genbank (ADIS-MP12 021)
JOURNAL
PUBMED 12004
Contact: Bernd Weisshaar
Chair of Genome Research
Bielefeld University, Institute for Genome Research
Universitaetsstrasse 25, D-33594 Bielefeld, Germany
Tel: +49-521-106 6873
Fax: +49-521-106 6423
Email: bernd.weisshaar@uni-bielefeld.de
Insert length: 624 Std Error: 0.00
Plate: 9 row: J column: 3
Seq primer: T7R CTATACGACTCATATAGGA.
Location/Qualifiers
source 1..624
/organism="Brassica napus"
/mol_type="mRNA"
/cultiVar="Express 617"
/db_xref="GABI:111175"
/db_xref="taxon:3708"
/clone="MP12p102J0039Q"
/issue_type="flowers and flower buds"
/dev_stage="young flowers and flower buds"
/clone_1lb="ADIS-MP12 021"
/note="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; cDNA
library from Brassica napus, strain Express 617; RNA was
from young flowers and flower buds of two greenhouse
plants; library was made at the Max-Planck-Institute for
Plant Breeding Research, Cologne, Germany; cloning sites
SalI-NotI, primer sites and orientation:
T7-salI-CCACGCGTCCG-5err-cDNA-polyA-CC-NotI-Spe; made by
Marc Jakoby 09/2000; PI: Bernd Weisshaar. Sequence
submission managed by RZPD/GABI-Primary database:

ORIGIN http://gabi.rzpd.de"

Query Match 14.9%; Score 359.2; DB 8; Length 624;
Best Local Similarity 76.9%; Pred. No. 5.6e-90;
Matches 518; Conservative 0; Mismatches 93; Indels 63; Gaps 4;

QY 1303 AGAATCCAGTATTGTGAGTTGTTTGGAAATTCGAATCGTGTATGACATATGAT 1362
DB 2 AGGAATCCAGTATTGTGACCTTGTCTGGAAATTCGAATCGTGTATGACATATGAC 61
QY 1363 CTCCTCGAGCTATGCAATTTGTTGAAACCTGTTGGCAGGGTGTCTTTCTAGGTTT 1422
DB 62 CTCCTCGAGCTATGCAATTTGTTGAAACCTGTTGGCAGGGTGTCTTTCTAGGTTT 121
QY 1423 AGAGACACCAAGATTTAAATTTAACTCGGGGATCTACATATGATGATCTTATGTTT 1482
DB 122 AGAGATCCCAAGATTTAAATTTAACTCGGGGATCTACATATGATGATCTTATGTTT 181
QY 1483 AGTACTTGGAAAAGTGGAGGTAGTTCAAGGTTCTCTTTAGCTGCTGCAACTATG 1542
DB 182 AGTACTTGGAAAAGTGGAGGTAGTTCAAGGTTCTCTTTAGCTGCTGCAACTATG 241
QY 1543 GGAAGATTGGAGCCGAGATGTGAAAGCTATGCTATGCAAGGCACTGCAAAAGTTT 1602
DB 242 GGAAGATTGGAGCCGAGATGTGAAAGCTATGCTATGCAAGGCACTGCAAAAGTTT 301
QY 1603 CCTTCCCGTATACAGATAGAAATCGGCTGAACCCAGATGTGCAAGACAGTTT 1662
DB 302 CCTTCCCGTATACAGATAGAAATCGGCTGAACCCAGATGTGCAAGACAGTTT 337
QY 1663 AGTGTAGATCTGTGTGTGTAACAATGAGCCGTGA---TGATGAGCTGTGTCTTAT 1719
DB 338 ACTGTAGATCATGTGTGTGTAACAATGAGCAATGATGTGATGAGCTGTGTATAGT 397
QY 1720 GGAAGATCTGTAAAGCCCTGTGAAAATTGAACTAATGATTAATGCAATTGAGCTGG 1779
DB 398 GGAAGATCTGTAAAGCCCTGTGAAAATTGAACTAATGATTAATGCAATTGAGCTGG 431
QY 1780 GTCTCAGAGATGAGGTGATGAAACTACTGTGTAATGCTGTTGCTGATATGTTAA 1839
DB 432 ----CAAGCTAGAGGTGATGAAACTACTGTGTAATGCTGTTGCTGATATGTTAA 487
QY 1840 GAGGCAAGTGTGAAGATCTTACGCTGTGTGTGGAATTTGCACTGTTCAAGC 1899
DB 488 GAGGCAAGTGTGAAGATCTTACGCTGTGTGTGGAATTTGCACTGTTCAAGC 547
QY 1900 CAGAAGTATTTCTTAAAA-----GCACTCATCTTTTCAAGCAAGATATGTTCT 1953
DB 548 CTGAGGTATATGTTCTATTAAAGCGCGCAGCTCATGTTGCAACGCAAGATATGACTTCC 607
QY 1954 TCTATGATCTGA 1967
DB 608 TCTACGCACTGA 621

RESULT 11
LOCUS CO075595
DEFINITION GR_Ea36G17.r GR_Ea Goeypium raimondii cDNA clone GR_Ea36G17.3',
mRNA sequence.
ACCESSION CO075595
VERSION CO075595.1 GI:48745076
KEYWORDS EST.
SOURCE Goeypium raimondii
ORGANISM Goeypium raimondii
Bukeriyota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Malvales; Malvaceae; Malvoideae; Goeypium.
1 (bases 1 to 769)
Kim,H., Yu,Y., Kudrna,D., Hatfield,J., Stum,D., Mueller,C.,
Udall,J.A., Rapp,R.A., Wendel,J.F., Rao,K., Soderlund,C. and
Wing,R.A.

TITLE Global assembly of Cotton ESTs
JOURNAL Unpublished (2004)
COMMENT Contact: Rod A. Wing
Arizona Genomics Institute
The University of Arizona
Forbes Building Room 303, Tucson, AZ, 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: rwing@genome.arizona.edu
Plate: 36 row: G column: 17.
FEATURES
source
1..769
location/Qualifiers
/organism="Goeypium raimondii"
/mol_type="mRNA"
/db_xref="taxon:29730"
/clone="GR_Ea36G17"
/tissue_type="whole seedlings"
/dev_stage="first true leaves"
/lab_host="DH10B"
/clone_id="GR_Ea"
/note="Vector: pCMVSPORT-6.1; Site 1: NotI; Site 2:
EcoRV; Library made by Invitrogen with RNA supplied by
Wendle lab. Directional cloned into NotI-EV. Colonies
plated/picked by AGI. More glycerol clones held in -80."

ORIGIN

Query Match 14.3%; Score 344.2; DB 7; Length 769;
Best Local Similarity 68.9%; Pred. No. 1.1e-85;
Matches 491; Conservative 0; Mismatches 213; Indels 9; Gaps 1;

QY 1 ATGGAAGCTGTGAGTCAAGTGGCAATTTGCTCTCCCAATTCGAATTCGCAATGCA 60
DB 54 ATGGAATCTTTGAGGACATATGATGATCGGGCTGTGCACTCTTCACTTATTC 113
QY 61 CCGGGAAGCAAAAGCTCCGAGTGAAGCCACAACACTTACATATGTGCTCCGCGAC 120
DB 114 CACCTCTTAAACCTTCAGACTCCACACCGTTCACTACCTGCTGCTGCGCAGT 173
QY 121 AAATGGGCGGACCGCTCTCTCCGACTTCAATTT-----CACCTCGATTCTCC 171
DB 174 AAATGGGCGGACCGCTCTCTCCGACTTCAATTT-----CACCTCGATTCTCC 233
QY 172 TCTCTCTCTTCCGCAACCGCAACCAACGCACTCTGCTCTTCCGCAACATCTAT 231
DB 224 TCTCTCTCTCTCTCTCAACCGCAACCTTCTCCCTTACCTGCTGCTGCTGCTG 293
QY 232 GATGCTCCGGAACGCAAGCTCCCACTCCCAATGATTTCTACAGATTTTGAAGCTAA 291
DB 294 TCTCTCTCCGGAACGCAAGCTCCCACTCCCAATGATTTCTACAGATTTTGAAGCTAA 353
QY 292 ACACATTTCTTAACCGATGAATGAAAGAGCATTCGAAGCTAGGTTTCGAACCGCG 351
DB 354 ACTCATTTCTTAAGGTGATGAAATGAAAGAGCTATGAAAGAGGTTTCGAACCGCGCT 413
QY 352 CAATTCGTTTCAAGCAAGAGCTTTTATCAGCCGAGACAGATTTCTTCAAGCTGCTGC 411
DB 414 CAATTCGTTTCAAGCAAGAGCTTTTATCAGCCGAGACAGATTTCTTCAAGCTGCTGC 473
QY 412 GAACTCTGTCTTATCTCGGTCTTGAAGAGATGAAAGTCTTCTTGAATGATA 471
DB 474 GAAACCTTATTTAAACCTCGGCTTGAAGAGATTTTCAACCAAGGCTTGTGACATAG 533
QY 472 GAACTCAGTCACTGATGATCTTCTTGAAGATTTCTTGAAGGCTCTGATGATATG 531
DB 534 CTGACACTATCATCTCAAGTCCCTCGGAGCAAGGTTCTTGAAGATTTGCTGCTG 593
QY 532 CAAGAAGGTGTGAGACTGAGATGTTCTTGGGTTGAGGCTCTGCTTAAAGAGAG 591
DB 594 CAAGAAGGTGTGAGACTGAGATGTTCTTGGGTTGAGGCTCTGCTTAAAGAGAGAG 653
QY 592 TTGCTAAGTGTGTTAAGCAAGATGTTTAAATGAGGCTTGTGCTTCTGATGTC 651
DB 654 CTGCCAAGGCAATTTAAACAGAGATGCTTGTGCAATGCGCTGCTTATGATGACTTG 713

QY 652 TCAGAGGATGCTATGATGATTCACCTGATTTTATTACTGTTAGATT 704
|||||
Db 714 TCAGAGGATGCTATGATGATTCACCTGATTTTATTACTGTTAGATT 766

RESULT 12
CX543653 741 bp mRNA linear EST 12-JAN-2005
LOCUS UCRPT01_5 008 C05 T3 Poncirus trifoliata CTV-challenged cDNA
DEFINITION library - UCRPT01-UCR2 Poncirus trifoliata cDNA clone
UCRPT01_008_T3_C05, mRNA sequence.
CX543653
CX543653.1 GI:57570678
EST.
Poncirus trifoliata
Poncirus trifoliata
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Sapindales; Rutaceae; Poncirus.
1 (bases 1 to 741)
Roose, M.L., Ye, X., Federici, C.F., Close, T.J., Fenton, R.D.,
Wanamaker, S., Choi, Y. and Kingan, T.
Development of EST Resources and New Genetic Markers for California
Citrus - Poncirus trifoliata CTV-challenged phloem - UCRPT01-UCR2
Unpublished (2004)
Contact: Mikeal Roose
Department of Botany & Plant Sciences, University of California
Riverside, CA 92521-0124, USA
Tel: 9097874137
Fax: 9097874437
Email: mikeal.roose@ucr.edu
Seq primer: T3.
Location/Qualifiers
1..741
/organism="Poncirus trifoliata"
/mol_type="mRNA"
/cultivar="Pomeroy OP"
/db_xref="taxon:37690"
/clone="UCRPT01_008_T3_C05"
/tissue_type="bark [with phloem]"
/dev_stage="10 - 30 cm shoots"
/lab_host="E. coli TUC121"
/clone_lib="Poncirus trifoliata CTV-challenged cDNA
library - UCRPT01-UCR2"
/note="Vector: Lambda Uni-ZAP XR, excised phagemid;
Site 1: EcoRI; Site 2: XhoI. Plants were grown in the
greenhouse at University of California, Riverside. The
screen was an open-pollinated (very probably selfed)
seedling of Poncirus trifoliata cv Pomeroy that was
selected as homozygous for the CTV resistance gene. The
rootstock was sweet orange infected with citrus tristeza
virus (CTV) isolate T514 over 1 year before sampling (CTV
infects sweet orange, but not genotypes carrying the CTV
resistance gene. Shoots 10-30 cm long were harvested in
October 2000, and the green phloem (bark) was removed and
frozen quickly in dry ice. Total RNA was extracted using
Trizol reagent (Gibco). Poly(A) RNA was purified, a cDNA
library was made, and 0.5 million primary lambda cDNA
clones were in vivo excised to give a population of
phagescript SK(-) phagemids. All steps to this point were
performed in the M. Roose lab at the University of
California, Riverside by X. Ye. Phagemids were plated,
plasmid DNA purified, cDNA clones archived, and DNA
sequences determined bi-directionally using an ABI730 at
the University of California Riverside Institute of
Integrative Genome Biology Genomics Core Instrumentation
Facility, (Choi, Kingan). Chromatogram files were
downloaded by FTP by Close, then processed by Wanmaker
(Close lab) using the Harvest pipeline
(http://harvest.ucr.edu) to remove vector and cloning
oligo sequences and various contaminants, and to trim to a
high quality region. Sequences that retained a phred 17

region of at least 100 bases were assembled, then chimeras
were removed following manual inspection of assemblies
(Close, Roose, Wanmaker). Sequences that survived all
removal steps were submitted to Genbank."

ORIGIN

Query Match 13.9%; Score 335.4; DB 8; Length 741;
Best Local Similarity 73.3%; Pred. No. 3,4e-83;
Matches 445; Conservative 0; Mismatches 156; Indels 6; Gaps 1;

QY 98 CTAAACATATCTGCTCGCGCAGAAATGGGCGACGCTCTTCTCCGACTGAATTCA 157
|||
Db 138 CCACACATCTCTCTCTCGCGCAAAATGGGCGACGCTCTTCTCCGACTGAATTCA 197
158 CCTCCGATTCCT 217
|||
Db 198 CCACGCGCGCAAAATCT 257
218 TGCCACCATCTATTGATGCTGTCGGAAGCGACGCTCCCATCCCATTTGTTACCAAG 277
Db 258 CCCCTCC-----TACTCCACCAACCGCACGATCATCCCATCCCATCTTATCAGG 311
278 TATTAGAGCTCAACACATTTCTTACCGATGGAATCAAGAGCAATTGAGCTAGG 337
|||
Db 312 CGTGGAGGCGAGACTCATTTCTTGGAGATGAATAGAGGCTTATGAGCTAGGA 371
338 TTTCGAACCGCGCAATTCGCTTTCAGCGACGCTTTAATCAGCCGAGACAGATTTC 397
|||
Db 372 TTTCATAACACATCCATAGAGGCTTATGAGCTTATGATCAGCGATGACAGATTTC 431
398 TTCAAGCTGCTTGGAAACTCTGCTTATCTCTGCTTGAAGAGATCAATGAGAGTTC 457
|||
Db 432 TTCAAGCTGCTTGGAAACTCTGCTTATCTCTGCTTGAAGAGATCAATGAGAGTTC 491
458 TTCTTGATGATGAAGAGATCACTATCACTGATGCTTCTGGGATPAAGTTCCTGAGG 517
|||
Db 492 TCGCTGATGACATCCGATACATCTCTGATGCTTCTGGAGACAGATTCCTGAG 551
518 CTCTCTGATGATGCAAGAGATGAGTGAAGCTGATGATGCTTCTGGGATGAGAGCTTC 577
|||
Db 552 CATTCGCTTGTGTTCAAGAGATGAGGAGAGCGAGATGATGAGATGAGAGATTC 611
578 TGCTTAAGAGAGATGCTCTTAAGCTTGAAGCAATGCTGTTTATGATGAGCTTCG 637
|||
Db 612 TGTTGAAGAGAGACTTCCCAAGCTTTCGAAGCAAGATGTTGCTTGGCGATGCACTTG 671
638 CGTTTCGATGCTGAGGATGCTTATGATGATGATGATGATGATGATGATGATGATG 697
|||
Db 672 CTTATGTTGACATTTCCAGGAGATGATGATGATGATGATGATGATGATGATGATG 731
QY 698 ATGAGTT 704
|||
Db 732 GTGAGAT 738

RESULT 13
BU046755 631 bp mRNA linear EST 26-AUG-2002
LOCUS PP_LEA0027104f Peach developing fruit mesocarp Prunus persica cDNA
DEFINITION clone PP_LEA0027104f, mRNA sequence.
BU046755
BU046755.1 GI:22486832
EST.
Prunus persica (peach)
Prunus persica
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Rosales; Rosaceae; Amygdaloideae; Prunus.
1 (bases 1 to 631)
Callahan, A., Palmer, M., Main, D., Wing, R. and Abbott, A.
Peach Model Genome for Rosaceae
Unpublished (2002)
Contact: Abbott, A.

Qy	479	CAGCATCAGTAGTTCCTTGGGGTAAGGTTCCCTGGGGGCTCTCGTGTATTGCAAG	538
Db	495	CCATCATCCTCAAGTTCCTTGGGATTAAGTTCCTGGAGCTCTGTGCTGTGCAAGAG	554
Qy	539	GTCGTGACATGAGATAGTTCCTTCGGGGTGTGAGGCTCTGCTTTAAGAGAGTTGCCTA	598
Db	555	CTGGGAAGACTAGAGCTGGTCTTCTTCAAAATTGGGGAGAGTTTCTTAAGAGAGGCTGCCCA	614
Qy	599	AGTCGTTAAACAAGATGTCGTTTATGTTATGCGGCTTGCCGTTCTCGATGTCGAGGG	658
Db	615	AGTGTTCACCAAGATGTCGTTTGGTCATGCGACCTTGTTATGTTGACATGTCGAGGG	674
Qy	659	ATGCTATGGCATTGGATCCACTGATTTT	687
Db	675	ATGCATAGGATTTGTCCCGCCTGATTTT	703

RESULT 15
BP785511/c

LOCUS	415 bp	mrna	linear	EST 10-FEB-2005
BP785511				
DEFINITION	BP785511	RAFL7	Arabidopsis thaliana	CDNA clone RAFL07-95-D11 3'
DESCRIPTION	mrna sequence.			

ACCESSION
VERSION

SOURCE ORGANISM	Arabidopsis thaliana (thale cress)	Arabidopsis thaliana
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15
16	16	16
17	17	17
18	18	18
19	19	19
20	20	20
21	21	21
22	22	22
23	23	23
24	24	24
25	25	25
26	26	26
27	27	27
28	28	28
29	29	29
30	30	30
31	31	31
32	32	32
33	33	33
34	34	34
35	35	35
36	36	36
37	37	37
38	38	38
39	39	39
40	40	40
41	41	41
42	42	42
43	43	43
44	44	44
45	45	45
46	46	46
47	47	47
48	48	48
49	49	49
50	50	50
51	51	51
52	52	52
53	53	53
54	54	54
55	55	55
56	56	56
57	57	57
58	58	58
59	59	59
60	60	60
61	61	61
62	62	62
63	63	63
64	64	64
65	65	65
66	66	66
67	67	67
68	68	68
69	69	69
70	70	70
71	71	71
72	72	72
73	73	73
74	74	74
75	75	75
76	76	76
77	77	77
78	78	78
79	79	79
80	80	80
81	81	81
82	82	82
83	83	83
84	84	84
85	85	85
86	86	86
87	87	87
88	88	88
89	89	89
90	90	90
91	91	91
92	92	92
93	93	93
94	94	94
95	95	95
96	96	96
97	97	97
98	98	98
99	99	99
100	100	100

REFERENCE

AUTHORS

TITLE
JOURNAL
COMMENT

Seki, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M., Akiyama, K., Iida, K., Enju, A., Sakurai, T., Arikawa, T., Carrinón, P., Fukuda, S., Iida, J., Kawai, J., Sasaki, D., Shiraki, T., Hayashizaki, Y. and Shimozaki, K.
Large-Scale Analysis of RIKEN Arabidopsis Full-length cDNAs Unpublished (2005)
Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyada, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-4350
Email: mseki@rsc.riken.go.jp
An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998, 2002). This clone is in a modified plasmid vector
Please visit our web site (<http://pfweb.gsc.riken.jp> and <http://large.gsc.riken.jp>) for further details.
reversed clone.

FEATURES	Location/Qualifiers
source	1. .415

```

1. .415
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAFL07-95-D11"
/dev_stage="rosette plants"
/lab_host="MDH08"
/clone_lib="RAFL7"
/note="Site 1: BamHI, Site 2: SalI; subjected to
cold-treated (1, 2, 5, 10, 24 hr)"

```

ORIGIN

Query Match 13.3%; Score 321; DB 3; Length 415;
Best Local Similarity 100.0%; Pred. No. 3.4e-79;
Matches 321; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2086 CCTGATCACC GCATAGAATGTTACCAGAGTTTGGATGGCGAATGCTGAAGATTGG 2145

Db 415 CCTGATCACC GCATAGAA TGTAC CAGAGTTT GGATGGCG AATGCTGA AGATTTCG 356

2146 ACTGACAGAGCAGCTGAACCTGCGCAGCTTGGGTTGGTTATGATTATACACTGTTGAA 2205

Db	355	ACTGACAGAGCAGCTGAAA	CTGCGCAGCTGGGTTTATGATTATACACTGTTGAAA	2366
Oy	2206	CTATCTGTTGACAGTGTG	CACTCTCAGACAAATGGAAACCGTGCTGTGGGAAAGCACT	2266
Db	295	CTATCTGTTGACAGTGTGACAGTCTCAGACAGATGGAACCGTGCTGTGGTGGAGCACT	2366	
Oy	2266	CTGAGAGAGTCTGCTTGTCTATCTGATTTGGTTCA	TCAGAAAACAATGCTACTGATGTC	2222
Db	235	CTGGAAGAGTCTGCTTGTCTATCTGATTTGGTTCA	TCAGAAAACAATGCTACTGATGTC	1766
Oy	2326	AGAACTTACACAACAAGATACAAATTTTGTGTC	CAAGTCAGGGTGGAAAAATCACTGAA	2388
Db	175	AGAACTTACACAACAAGATACGAATTTTCTGTGTC	CAAGTCAGGGTGGAAAAATCACTGAA	1166
Oy	2386	GGCTCTGTTCTTGCAATCAA	2406	
Db	115	GGCTCTGTTCTTGCAATCAA	95	

Search completed: December 11, 2005, 10:07:47
Job time : 9158.75 secs

```
QY      2386 GGCTCTGTTCCTGCATCATAA 2406
          |||||
Db      115  GGCTCTGTTCCTGCATCATAA 95
```